

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
3 June 2004 (03.06.2004)

PCT

(10) International Publication Number
WO 2004/046340 A2

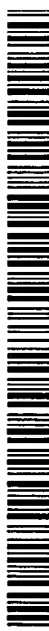
- (51) International Patent Classification⁷: C12N
- (21) International Application Number: PCT/US2003/037047
- (22) International Filing Date:
14 November 2003 (14.11.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/426,095 14 November 2002 (14.11.2002) US
- (71) Applicant (for all designated States except US): GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): KRUMMEN, Lynne [US/US]; 3030 22nd Avenue, San Francisco, CA 94132 (US). SHEN, Amy [US/US]; 1828 Parrott Drive, San Mateo, CA 94402 (US).
- (74) Agents: STEFFES, David et al.; Sidley Austin Brown & Wood LLP, 1501 K Street, N.W., Washington, D.C. 20005 (US).

- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 2004/046340 A2

(54) Title: INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-EXPRESSING PRODUCTION CELL LINES

(57) Abstract: This invention relates to a DNA construct, methods of selecting for high-expressing host cells, a method of producing a protein of interest in high yields and a method of producing eukaryotic cells having multiple copies of a sequence encoding a protein of interest. In one method, stable clones capable of producing a high level of a product of interest are generated from one step of a direct selection immediately after transfection.

INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-
EXPRESSING PRODUCTION CELL LINES

This application claims priority under 35 U.S.C. § 119(e) from U.S. provisional application serial no. 60/426,095, filed November 14, 2002, which is herein incorporated by reference.

BACKGROUND OF THE INVENTION

Field of the Invention

This invention relates to a DNA construct, a method of selecting for high-expressing host cells, a method of producing a protein of interest in high yields and a method of producing eukaryotic cells having multiple copies of a sequence encoding a protein of interest.

Description of Background and Related Art

The discovery of methods for introducing DNA into living host cells in a functional form has provided the key to understanding many fundamental biological processes, and has made possible the production of important proteins and other molecules in commercially useful quantities.

Despite the general success of such gene transfer methods, several common problems exist that may limit the efficiency with which a gene encoding a desired protein can be introduced into and expressed in a host cell. One problem is knowing when the gene has been successfully transferred into recipient cells. A second problem is distinguishing between those cells that contain the gene and those that have survived the transfer procedures but do not contain the gene. A third problem is identifying and isolating those cells that contain the gene and that are expressing high levels of the protein encoded by the gene.

In general, the known methods for introducing genes into eukaryotic cells tend to be highly inefficient. Of the cells in a given culture, only a small proportion take up and express exogenously added DNA, and an even smaller proportion stably maintain that DNA.

Identification of those cells that have incorporated a product gene encoding a desired protein typically is achieved by introducing into the same cells another gene, commonly referred to

as a selectable gene, that encodes a selectable marker. A selectable marker is a protein that is necessary for the growth or survival of a host cell under the particular culture conditions chosen, such as an enzyme that confers resistance to an antibiotic or other drug, or an enzyme that compensates for a metabolic or catabolic defect in the host cell. For example, selectable genes commonly used with eukaryotic cells include the genes for aminoglycoside phosphotransferase (APH), hygromycin phosphotransferase (hyg), dihydrofolate reductase (DHFR), thymidine kinase (tk), neomycin resistance, puromycin resistance, glutamine synthetase, and asparagine synthetase.

The method of identifying a host cell that has incorporated one gene on the basis of expression by the host cell of a second incorporated gene encoding a selectable marker is referred to as cotransfection (or cotransfection). In that method, a gene encoding a desired polypeptide and a selection gene typically are introduced into the host cell simultaneously. In this case of simultaneous cotransfection, the gene encoding the desired polypeptide and the selectable gene may be present on a single DNA molecule or on separate DNA molecules prior to being introduced into the host cells. Wigler *et al.*, Cell, 16:777 (1979). Cells that have incorporated the gene encoding the desired polypeptide then are identified or isolated by culturing the cells under conditions that preferentially allow for the growth or survival of those cells that synthesize the selectable marker encoded by the selectable gene.

The level of expression of a gene introduced into a eukaryotic host cell depends on multiple factors, including gene copy number, efficiency of transcription, messenger RNA (mRNA) processing, stability, and translation efficiency. Accordingly, high level expression of a desired polypeptide typically will involve optimizing one or more of those factors.

For example, the level of protein production may be increased by covalently joining the coding sequence of the gene to a "strong" promoter or enhancer that will give high levels of transcription. Promoters and enhancers are nucleotide sequences that interact specifically with proteins in a host cell that are involved in transcription. Kriegler, Meth. Enzymol., 185:512 (1990); Maniatis *et al.*, Science, 236:1237 (1987). Promoters are located upstream of the coding sequence of a gene and facilitate transcription of the gene by RNA polymerase. Among the eukaryotic promoters that have been identified as strong promoters for high-level expression are the SV40 early promoter, adenovirus major late promoter, mouse metallothionein-I promoter, Rous sarcoma virus long terminal repeat, and human cytomegalovirus immediate early promoter (CMV).

Enhancers stimulate transcription from a linked promoter. Unlike promoters, enhancers are active when placed downstream from the transcription initiation site or at considerable distances from the promoter, although in practice enhancers may overlap physically and functionally with promoters. For example, all of the strong promoters listed above also contain strong enhancers. Bendig, Genetic Engineering, 7:91 (Academic Press, 1988).

The level of protein production also may be increased by increasing the gene copy number in the host cell. One method for obtaining high gene copy number is to directly introduce into the host cell multiple copies of the gene, for example, by using a large molar excess of the product gene relative to the selectable gene during cotransfection. Kaufman, Meth. Enzymol., 185:537 (1990). With this method, however, only a small proportion of the cotransfected cells will contain the product gene at high copy number. Furthermore, because no generally applicable, convenient method exists for distinguishing such cells from the majority of cells that contain fewer copies of the product gene, laborious and time-consuming screening methods typically are required to identify the desired high-copy number transfectants.

Another method for obtaining high gene copy number involves cloning the gene in a vector that is capable of replicating autonomously in the host cell. Examples of such vectors include mammalian expression vectors derived from Epstein-Barr virus or bovine papilloma virus, and yeast 2-micron plasmid vectors. Stephens & Hentschel, Biochem. J., 248:1 (1987); Yates *et al.*, Nature, 313:812 (1985); Beggs, Genetic Engineering, 2:175 (Academic Press, 1981).

Yet another method for obtaining high gene copy number involves gene amplification in the host cell. Gene amplification occurs naturally in eukaryotic cells at a relatively low frequency. Schimke, J. Biol. Chem., 263:5989 (1988). However, gene amplification also may be induced, or at least selected for, by exposing host cells to appropriate selective pressure. For example, in many cases it is possible to introduce a product gene together with an amplifiable gene into a host cell and subsequently select for amplification of the marker gene by exposing the cotransfected cells to sequentially increasing concentrations of a selective agent. Typically the product gene will be coamplified with the marker gene under such conditions.

The most widely used amplifiable gene for that purpose is a DHFR gene, which encodes a dihydrofolate reductase enzyme. The selection conditions used in conjunction with a DHFR gene are the absence of glycine, hypoxanthine and thymidine (GHT) with or without the presence of methotrexate (Mtx). A host cell is cotransfected with a product gene encoding a desired protein

and a DHFR gene, and transfectants are identified by first culturing the cells in GHT-free culture medium that may contains Mtx. A suitable host cell when a wild-type DHFR gene is used is the Chinese Hamster Ovary (CHO) cell line deficient in DHFR activity, prepared and propagated as described by Urlaub & Chasin, Proc. Nat. Acad. Sci. USA, 77:4216 (1980). The transfected cells then are exposed to successively higher amounts of Mtx. This leads to the synthesis of multiple copies of the DHFR gene, and concomitantly, multiple copies of the product gene. Schimke, J. Biol. Chem., 263:5989 (1988); Axel *et al.*, U.S. Patent No. 4,399,216; Axel *et al.*, U.S. Patent No. 4,634,665. Other references directed to co-transfection of a gene together with a genetic marker that allows for selection and subsequent amplification include Kaufman in Genetic Engineering, ed. J. Setlow (Plenum Press, New York), Vol. 9 (1987); Kaufman and Sharp, J. Mol. Biol., 159:601 (1982); Ringold *et al.*, J. Mol. Appl. Genet., 1:165-175 (1981); Kaufman *et al.*, Mol. Cell Biol., 5:1750-1759 (1985); Kaetzel and Nilson, J. Biol. Chem., 263:6244-6251 (1988); Hung *et al.*, Proc. Natl. Acad. Sci. USA, 83:261-264 (1986); Kaufman *et al.*, EMBO J., 6:87-93 (1987); Johnston and Kucey, Science, 242:1551-1554 (1988); Urlaub *et al.*, Cell, 33:405-412 (1983).

To extend the DHFR amplification method to other cell types, a mutant DHFR gene that encodes a protein with reduced sensitivity to methotrexate may be used in conjunction with host cells that contain normal numbers of an endogenous wild-type DHFR gene. Simonsen and Levinson, Proc. Natl. Acad. Sci. USA, 80:2495 (1983); Wigler *et al.*, Proc. Natl. Acad. Sci. USA, 77:3567-3570 (1980); Haber and Schimke, Somatic Cell Genetics, 8:499-508 (1982).

Alternatively, host cells may be co-transfected with the product gene, a DHFR gene, and a dominant selectable gene, such as a neo^r gene. Kim and Wold, Cell, 42:129 (1985); Capon *et al.*, U.S. Pat. No. 4,965,199. Transfectants are identified by first culturing the cells in culture medium containing neomycin (or the related drug G418), and the transfectants so identified then are selected for amplification of the DHFR gene and the product gene by exposure to successively increasing amounts of Mtx.

As will be appreciated from this discussion, the selection of recombinant host cells that express high levels of a desired protein generally is a multi-step process. In the first step, initial transfectants are selected that have incorporated the product gene and the selectable gene. In subsequent steps, the initial transfectants are subject to further selection for high-level expression of the selectable gene and then random screening for high-level expression of the product gene. To identify cells expressing high levels of the desired protein, typically one must screen large numbers

of transfectants. The majority of transfectants produce less than maximal levels of the desired protein. Further, Mtx resistance in DHFR transformants is at least partially conferred by varying degrees of gene amplification. Schimke, Cell, 37:705-713 (1984). The inadequacies of co-expression of the non-selected gene have been reported by Wold *et al.*, Proc. Natl. Acad. Sci. USA, 76:5684-5688 (1979). Instability of the amplified DNA is reported by Kaufman and Schimke, Mol. Cell Biol., 1:1069-1076 (1981); Haber and Schimke, Cell, 26:355-362 (1981); and Fedespiel *et al.*, J. Biol. Chem., 259:9127-9140 (1984).

Several methods have been described for directly selecting such recombinant host cells in a single step. One strategy involves co-transfecting host cells with a product gene and a DHFR gene, and selecting those cells that express high levels of DHFR by directly culturing in medium containing a high concentration of Mtx. Many of the cells selected in that manner also express the co-transfected product gene at high levels Page and Sydenham, Bio/Technology, 9:64 (1991). This method for single-step selection suffers from certain drawbacks that limit its usefulness. High-expressing cells obtained by direct culturing in medium containing a high level of a selection agent may have poor growth and stability characteristics, thus limiting their usefulness for long-term production processes Page and Snyderman, Bio/Technology, 9:64 (1991). Single-step selection for high-level resistance to Mtx may produce cells with an altered, Mtx-resistant DHFR enzyme, or cells that have altered Mtx transport properties, rather than cells containing amplified genes. Haber *et al.*, J. Biol. Chem., 256:9501 (1981); Assaraf and Schimke, Proc. Natl. Acad. Sci. USA, 84:7154 (1987).

Another method involves the use of polycistronic mRNA expression vectors containing a product gene at the 5' end of the transcribed region and a selectable gene at the 3' end. Because translation of the selectable gene at the 3' end of the polycistronic mRNA is inefficient, such vectors exhibit preferential translation of the product gene and require high levels of polycistronic mRNA to survive selection. Kaufman, Meth. Enzymol., 185:487 (1990); Kaufman, Meth. Enzymol., 185:537 (1990); Kaufman *et al.*, EMBO J., 6:187 (1987). Accordingly, cells expressing high levels of the desired protein product may be obtained in a single step by culturing the initial transfectants in medium containing a selection agent appropriate for use with the particular selectable gene. However, the utility of these vectors is variable because of the unpredictable influence of the upstream product reading frame on selectable marker translation and because the upstream reading frame sometimes becomes deleted during methotrexate amplification (Kaufman

et al., J. Mol. Biol., **159**:601-621 (1982); Levinson, Methods in Enzymology, San Diego: Academic Press, Inc. (1990)). Later vectors incorporated an internal translation initiation site derived from members of the picornavirus family which is positioned between the product gene and the selectable gene (Pelletier *et al.*, Nature, **334**:320 (1988); Jang *et al.*, J. Virol., **63**:1651 (1989)).

A third method for single-step selection involves use of a DNA construct with a selectable gene containing an intron within which is located a gene encoding the protein of interest. See U.S. Patent No. 5,043,270 and Abrams *et al.*, J. Biol. Chem., **264**(24): 14016-14021 (1989). In yet another single-step selection method, host cells are co-transfected with an intron-modified selectable gene and a gene encoding the protein of interest. See WO 92/17566, published October 15, 1992. The intron-modified gene is prepared by inserting into the transcribed region of a selectable gene an intron of such length that the intron is correctly spliced from the corresponding mRNA precursor at low efficiency, so that the amount of selectable marker produced from the intron-modified selectable gene is substantially less than that produced from the starting selectable gene. These vectors help to insure the integrity of the integrated DNA construct, but transcriptional linkage is not achieved as selectable gene and the protein gene are driven by separate promoters.

Other mammalian expression vectors that have single transcription units have been described. Retroviral vectors have been constructed (Cepko *et al.*, Cell, **37**:1053-1062 (1984)) in which a cDNA is inserted between the endogenous Moloney murine leukemia virus (M-MuLV) splice donor and splice acceptor sites which are followed by a neomycin resistance gene. This vector has been used to express a variety of gene products following retroviral infection of several cell types.

A method for selecting recombinant host cells expressing high levels of a desired protein was previously described by the applicants in Lucas *et al.*, Nucleic Acid Research, **24**, No. 9: 1774-1779 and U.S. Patent No. 5,561,053. That method utilizes eukaryotic host cells harboring a DNA construct comprising a selectable gene (preferably an amplifiable gene) and a product gene provided 3' to the selectable gene. The selectable gene is positioned within an intron defined by a splice donor site and a splice acceptor site and the selectable gene and product gene are under the transcriptional control of a single transcriptional regulatory region. The splice donor site is generally an efficient splice donor site and thereby regulates expression of the product gene using the transcriptional regulatory region. The transfected cells are cultured so as to express the gene encoding the product in a selective medium which may contain an amplifying agent for sufficient

time to allow cells having multiple copies of the product gene, or cells with a single (or multiple) copy of the gene in a chromosomal loci with high transcriptional activity to be identified.

Other fusion expression constructs have been developed. For example, a fusion of green fluorescent protein with the Zeocin-resistance marker construct has been created. Bennet, R.P. *et al.*, Biotechniques, 24(3):478-82, 1998 March. Such constructs were used to allow visual screening and drug selection of transfected eukaryotic cells.

In another example, human prothrombin was overexpressed in transformed eukaryotic cells using a dominant bifunctional selection and amplification marker. Herlitschka, Sabine E. *et al.*, *Protein Expression and Purification*, 8, 358-364, 1996 July. In this reference the marker consisted of the murine wild-type dihydrofolate reductase cDNA and the *E. coli* hygromycin phosphotransferase gene fused in frame. The gene of interest is connected, upstream, by the EMCV untranslated region to the fusion marker gene, forming a dicistronic transcription unit.

With the state of the art in mind, it is one object of the present invention to increase the level of homogeneity with regard to expression levels of stable clones transfected with a product gene of interest, by expressing fused selectable markers (i.e. DHFR and puromycin) and a protein of interest from a single promoter.

It is another object to provide a method for selecting stable, recombinant host cells that express high levels of a desired protein product, which method is rapid and convenient to perform, and reduces the numbers of transfected cells which need to be screened. Furthermore, it is an object to allow high levels of single and multiple unit polypeptides to be rapidly generated from clones or pools of stable host cell transfectants.

It is an additional object to provide expression vectors which bias for active integration events (i.e. have an increased tendency to generate transformants wherein the DNA construct is inserted into a region of the genome of the host cell which results in high level expression of the product gene) and can accommodate a variety of product genes without the need for modification.

SUMMARY OF THE INVENTION

Accordingly, the present invention is directed to a DNA construct (DNA molecule) comprising a 5' transcriptional initiation site and a 3' transcriptional termination site, two selectable genes that have been fused into one open reading frame (preferably amplifiable genes) and a product gene provided 3' to the fused selectable genes, a transcriptional regulatory region

regulating transcription of both the fused selectable genes and the product gene, the fused selectable genes positioned within an intron defined by a splice donor site and a splice acceptor site. The splice donor site preferably comprises an effective splice donor sequence as herein defined and thereby regulates expression of the product gene using the transcriptional regulatory region.

In another embodiment, the invention provides a method for producing a product of interest comprising culturing a eukaryotic cell which has been transfected with the DNA construct described above, so as to express the product gene and recovering the product.

In a further embodiment, the invention provides a method for producing eukaryotic cells having multiple copies of the product gene comprising transfecting eukaryotic cells with the DNA construct described above (where the selectable fused genes are amplifiable genes), growing the cells in a selective medium comprising an amplifying agent(s) for a sufficient time for amplification to occur, and selecting cells having multiple copies of the product gene. After transfection of the host cells, most of the transfectants fail to exhibit the selectable phenotype characteristic of the protein encoded by either of the selectable genes, but surprisingly a small proportion of the transfectants do exhibit one or both of the selectable phenotype, and among those transfectants, the majority are found to express high levels of the desired product encoded by the product gene. Thus, the invention provides an improved method for the selection of recombinant host cells expressing high levels of a desired product, which method is useful with a wide variety of eukaryotic host cells and avoids the problems inherent in, and improves upon, existing cell selection technology.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates schematically the construction of the pSV.IPD. The gene for the protein of interest would be inserted at the polylinker site.

Figures 2-1 to 2-4 depict the nucleotide sequence of the pSV.IPUR plasmid used in constructing pSV.IPD (SEQ ID NO 1).

Figures 3-1 to 3-4 depict the nucleotide sequence of the pSV.ID plasmid used in constructing pSV.IPD (SEQ ID NO 2).

Figures 4-1 to 4-4 depict the nucleotide sequence of the pSV.IPD (SEQ ID NO 3).

Figure 5 illustrates schematically the plasmid, pSV.ID.VEGF, used as a control in Example 1.

Figure 6 illustrates schematically the plasmid, pSV.IPD.2C4, used in Example 1 (SEQ ID NO 4).

Figures 7-1 to 7-8 depict the nucleotide sequence of the pSV.IPD.2C4 plasmid used in Example 1.

Figure 8 depicts a FACS analysis of transiently transfected CHO cells with a GTP plasmid in 250ml spinner transfection. FACS analysis was performed 24 hours after transfection.

Figure 9 depicts the expression level of clones from traditional 10nM MTX selection. Cells were transfected with commercial transfection reagent and directly selected in 10 nM MTX. Individual clones were grown in a 96-well plate. Product accumulated for 6 days prior to ELISA.

Figures 10-1 and 10-2 depict the expression level of clones from 25 and 50 nM MTX direct selections, respectively, of SV40-based constructs derived from spinner transfection. The assay was performed the same as in Figure 9.

Figure 11 depicts the expression level of clones from 25 nM MTX direct selection of CMV-based construct derived from spinner transfection. The assay was performed the same as in Figure 9.

Figure 12 depicts the titer evaluation in Miniform. Samples were collected every day and submitted to an HPLC protein A assay for titer.

Figure 13-1 to 13-7 depict the nucleotide sequence of the pCMV.IPD.Heterologous polypeptide (HP) plasmid used in Example 3.

Figure 14-1 to 14-8 depicts the nucleotide sequence of the pSV40.IPD.HP plasmid used in Example 3.

Figure 15 illustrates schematically the plasmid, pCMV.IPD.HP, used in Example 3.

Figure 16 illustrates a time line and titer comparison between a traditional selection and direct selection method described in Example 3. Equivalent titers are indicated horizontally across the illustration. For example, the titers for a 200/300nM SV40-plasmid traditional selection, 100nM SV40-plasmid direct selection and 25nM CMV-plasmid direct selection are roughly equivalent.

DESCRIPTION OF THE PREFERRED EMBODIMENT

Definitions:

The "DNA construct" disclosed herein comprises a non-naturally occurring DNA molecule or chemical analog which can either be provided as an isolate or integrated in another DNA molecule *e.g.* in an expression vector or the chromosome of an eukaryotic host cell.

The term "selectable gene" as used herein refers to a DNA that encodes a selectable marker necessary for the growth or survival of a host cell under the particular cell culture conditions chosen. Accordingly, a host cell that is transformed with a selectable gene will be capable of growth or survival under certain cell culture conditions wherein a non-transfected host cell is not capable of growth or survival. Typically, a selectable gene will confer resistance to a drug or compensate for a metabolic or catabolic defect in the host cell. Examples of selectable genes are provided in the following table. See also Kaufman, Methods in Enzymology, 185: 537-566 (1990), for a review of these.

"Fused selectable genes" as used herein refers to a DNA that encodes at least two selectable markers in the same open reading frame and inserted into an intron sequence.

TABLE 1**Examples of Selectable Genes and their Selection Agents**

Selection Agent	Selectable Gene
Puromycin	Puromycin-N-acetyltransferase
Methotrexate	Dihydrofolate reductase
Cadmium	Metallothionein
PALA	CAD
Xyl-A-or adenosine and 2'-deoxycoformycin	Adenosine deaminase
Adenine, azaserine, and coformycin	Adenylate deaminase
6-Azaauridine, pyrazofuran	UMP Synthetase
Mycophenolic acid	IMP 5'-dehydrogenase
Mycophenolic acid with limiting xanthine	Xanthine-guanine phosphoribosyltransferase
Hypoxanthine, aminopterin, and thymidine (HAT)	Mutant HGPRTase or mutant thymidine kinase
5-Fluorodeoxyuridine	Thymidylate synthetase
Multiple drugs e.g. adriamycin, vincristine or colchicine	P-glycoprotein 170
Aphidicolin	Ribonucleotide reductase
Methionine sulfoximine	Glutamine synthetase

β -Aspartyl hydroxamate or Albizziin	Asparagine synthetase
Canavanine	Arginosuccinate synthetase
α -Difluoromethylornithine	Ornithine decarboxylase
Compactin	HMG-CoA reductase
Tunicamycin	N-Acetylglucosaminyl transferase
Borrelidin	Threonyl-tRNA synthetase
Ouabain	Na ⁺ K ⁺ -ATPase

The preferred selectable genes are amplifiable genes. As used herein, the term "amplifiable gene" refers to a gene which is amplified (*i.e.* additional copies of the gene are generated which survive in intrachromosomal or extrachromosomal form) under certain conditions. The amplifiable gene(s) usually encodes an enzyme (*i.e.* an amplifiable marker) which is required for growth of eukaryotic cells under those conditions. For example, the gene may encode DHFR which is amplified when a host cell transformed therewith is grown in Mtx. According to Kaufman, the selectable genes in Table 1 above can also be considered amplifiable genes. An example of a selectable gene which is generally not considered to be an amplifiable gene is the neomycin resistance gene (Cepko *et al.*, *supra*).

As used herein, "selective medium" refers to nutrient solution used for growing eukaryotic cells which have the selectable gene(s) and therefore is deficient in components supplied by the selectable gene or includes a "selection agent". Commercially available media based on formulations such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) are exemplary nutrient solutions. In addition, any of the media described in Ham and Wallace, Meth. Enz., 58:44 (1979), Barnes and Sato, Anal. Biochem., 102:255 (1980), U.S. Patent Nos. 4,767,704; 4,657,866; 4,927,762; or 4,560,655; WO 90/03430; WO 87/00195; U.S. Patent Re. 30,985; or U.S. Patent No. 5,122,469, the disclosures of all of which are incorporated herein by reference, may be used as culture media. Any of these media may be supplemented as necessary with hormones and/or other

growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleosides (such as adenosine and thymidine), antibiotics (such as GentamycinTM drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The preferred nutrient solution comprises fetal bovine serum.

The term "selection agent" refers to a substance that interferes with the growth or survival of a host cell possibly because the cell is deficient in a particular selectable gene. Examples of selection agents are presented in Table 1 above. The selection agent preferably comprises an "amplifying agent" which is defined for purposes herein as an agent for amplifying copies of the amplifiable gene or causing integration of multiple copies of the amplifiable gene into the genome, such as Mtx if the amplifiable gene is DHFR. See Table 1 for examples of amplifying agents.

As used herein, the terms "direct selection" or "direct culturing" means the first exposure to selective conditions either without MTX or GHT or with MTX, and production of a heterologous polypeptide in an amount of about 250mg/l, 400mg/l, 600mg/l or 800mg/l up to about 1g/l or more.

As used herein, the term "transcriptional initiation site" refers to the nucleic acid in the DNA construct corresponding to the first nucleic acid incorporated into the primary transcript, *i.e.*, the mRNA precursor, which site is generally provided at, or adjacent to, the 5' end of the DNA construct.

The term "transcriptional termination site" refers to a sequence of DNA, normally represented at the 3' end of the DNA construct, that causes RNA polymerase to terminate transcription.

As used herein, "transcriptional regulatory region" refers to a region of the DNA construct that regulates transcription of the selectable gene and the product gene. The transcriptional regulatory region normally refers to a promoter sequence (*i.e.* a region of DNA involved in binding of RNA polymerase to initiate transcription) which can be constitutive or inducible and, optionally, an enhancer (*i.e.* a *cis*-acting DNA element, usually from about 10-300 bp, that acts on a promoter to increase its transcription).

As used herein, "product gene" refers to DNA that encodes a desired protein or polypeptide product. Any product gene that is capable of expression in a host cell may be used, although the methods of the invention are particularly suited for obtaining high-level expression of a product gene that is not also a selectable or amplifiable gene. Accordingly, the protein or polypeptide encoded by a product gene typically will be one that is not necessary for the growth or survival of a host cell under the particular cell culture conditions chosen. For example, product genes suitably encode a peptide, or may encode a polypeptide sequence of amino acids for which the chain length is sufficient to produce higher levels of tertiary and/or quaternary structure.

Examples of bacterial polypeptides or proteins include, *e.g.*, alkaline phosphatase and β -lactamase. Examples of mammalian polypeptides or proteins include molecules such as renin; a growth hormone, including human growth hormone, and bovine growth hormone; growth hormone releasing factor; parathyroid hormone; thyroid stimulating hormone; lipoproteins; alpha-1-antitrypsin; insulin A-chain; insulin B-chain; proinsulin; follicle stimulating hormone; calcitonin; luteinizing hormone; glucagon; clotting factors such as factor VIIIc, factor IX, tissue factor, and von Willebrands factor; anti-clotting factors such as Protein C; atrial natriuretic factor; lung surfactant; a plasminogen activator, such as urokinase or human urine or tissue-type plasminogen activator (t-PA); bombesin; thrombin; hemopoietic growth factor; tumor necrosis factor-alpha and -beta; enkephalinase; RANTES (regulated on activation normally T-cell expressed and secreted); human macrophage inflammatory protein (MIP-1-alpha); a serum albumin such as human serum albumin; mullerian-inhibiting substance; relaxin A-chain; relaxin B-chain; prorelaxin; mouse gonadotropin-associated peptide; a microbial protein, such as beta-lactamase; DNase; inhibin; activin; vascular endothelial growth factor (VEGF); receptors for hormones or growth factors; integrin; protein A or D; rheumatoid factors; a neurotrophic factor such as bone-derived neurotrophic factor (BDNF), neurotrophin-3, -4, -5, or -6 (NT-3, NT-4, NT-5, or NT-6), or a nerve growth factor such as NGF- β ; platelet-derived growth factor (PDGF); fibroblast growth factor such as aFGF and bFGF; epidermal growth factor (EGF); transforming growth factor (TGF) such as TGF-alpha and TGF-beta, including TGF- β 1, TGF- β 2, TGF- β 3, TGF- β 4, or TGF- β 5; insulin-like growth factor-I and -II (IGF-I and IGF-II); des(1-3)-IGF-I (brain IGF-I), insulin-like growth factor binding proteins; CD proteins such as CD-3, CD-4, CD-8, and CD-19; erythropoietin; osteoinductive factors; immunotoxins; a bone morphogenetic protein (BMP); an interferon such as interferon-alpha, -beta, and -gamma; colony stimulating factors (CSFs), *e.g.*, M-CSF, GM-CSF,

and G-CSF; interleukins (ILs), e.g., IL-1 to IL-10; superoxide dismutase; T-cell receptors; surface membrane proteins; decay accelerating factor; viral antigen such as, for example, a portion of the AIDS envelope; transport proteins; homing receptors; addressins; regulatory proteins; antibodies; chimeric proteins such as immunoadhesins and fragments of any of the above-listed polypeptides.

The product gene preferably does not consist of an anti-sense sequence for inhibiting the expression of a gene present in the host. Preferred proteins herein are therapeutic proteins such as TGF- β , TGF- α , PDGF, EGF, FGF, IGF-I, DNase, plasminogen activators such as t-PA, clotting factors such as tissue factor and factor VIII, hormones such as relaxin and insulin, cytokines such as IFN- γ , chimeric proteins such as TNF receptor IgG immunoadhesin (TNFr-IgG) or antibodies such as anti-IgE. An example of an antibody that can be produced with the pSV.IDP plasmid (Figure 4) is anti-HER2 Neu antibody, 2C4, as provided in Example 1, *supra*.

The term "intron" as used herein refers to a nucleotide sequence present within the transcribed region of a gene or within a messenger RNA precursor, which nucleotide sequence is capable of being excised, or spliced, from the messenger RNA precursor by a host cell prior to translation. Introns suitable for use in the present invention are suitably prepared by any of several methods that are well known in the art, such as purification from a naturally occurring nucleic acid or *de novo* synthesis. The introns present in many naturally occurring eukaryotic genes have been identified and characterized. Mount, Nuc. Acids Res., **10**:459 (1982). Artificial introns comprising functional splice sites also have been described. Winey *et al.*, Mol. Cell Biol., **9**:329 (1989); Gattermann *et al.*, Mol. Cell Biol., **9**:1526 (1989). Introns may be obtained from naturally occurring nucleic acids, for example, by digestion of a naturally occurring nucleic acid with a suitable restriction endonuclease, or by PCR cloning using primers complementary to sequences at the 5' and 3' ends of the intron. Alternatively, introns of defined sequence and length may be prepared synthetically using various methods in organic chemistry. Narang *et al.*, Meth. Enzymol., **68**:90 (1979); Caruthers *et al.*, Meth. Enzymol., **154**:287 (1985); Froehler *et al.*, Nuc. Acids Res., **14**:5399 (1986).

As used herein "splice donor site" or "SD" refers to the DNA sequence immediately surrounding the exon-intron boundary at the 5' end of the intron, where the "exon" comprises the nucleic acid 5' to the intron. Many splice donor sites have been characterized and Ohshima *et al.*, J. Mol. Biol., **195**:247-259 (1987) provides a review of these. An "efficient splice donor sequence" refers to a nucleic acid sequence encoding a splice donor site wherein the efficiency of splicing of

messenger RNA precursors having the splice donor sequence is between about 80 to 99% and preferably 90 to 95% as determined by quantitative PCR. Examples of efficient splice donor sequences include the wild type (WT) splice donor sequence and the GAC:GTAAGT sequence of Example 3. Other efficient splice donor sequences can be readily selected using the techniques for measuring the efficiency of splicing disclosed herein.

The terms "PCR" and "polymerase chain reaction" as used herein refer to the *in vitro* amplification method described in US Patent No. 4,683,195 (issued July 28, 1987). In general, the PCR method involves repeated cycles of primer extension synthesis, using two DNA primers capable of hybridizing preferentially to a template nucleic acid comprising the nucleotide sequence to be amplified. The PCR method can be used to clone specific DNA sequences from total genomic DNA, cDNA transcribed from cellular RNA, viral or plasmid DNAs. Wang & Mark, in PCR Protocols, pp.70-75 (Academic Press, 1990); Scharf, in PCR Protocols, pp. 84-98; Kawasaki & Wang, in PCR Technology, pp. 89-97 (Stockton Press, 1989). Reverse transcription-polymerase chain reaction (RT-PCR) can be used to analyze RNA samples containing mixtures of spliced and unspliced mRNA transcripts. Fluorescently tagged primers designed to span the intron are used to amplify both spliced and unspliced targets. The resultant amplification products are then separated by gel electrophoresis and quantitated by measuring the fluorescent emission of the appropriate band(s). A comparison is made to determine the amount of spliced and unspliced transcripts present in the RNA sample.

One preferred splice donor sequence is a "consensus splice donor sequence". The nucleotide sequences surrounding intron splice sites, which sequences are evolutionarily highly conserved, are referred to as "consensus splice donor sequences". In the mRNAs of higher eukaryotes, the 5' splice site occurs within the consensus sequence AG:GUAAGU (wherein the colon denotes the site of cleavage and ligation). In the mRNAs of yeast, the 5' splice site is bounded by the consensus sequence :GUAUGU. Padgett, *et al.*, Ann. Rev. Biochem., 55:1119 (1986).

The expression "splice acceptor site" or "SA" refers to the sequence immediately surrounding the intron-exon boundary at the 3' end of the intron, where the "exon" comprises the nucleic acid 3' to the intron. Many splice acceptor sites have been characterized and Ohshima *et al.*, J. Mol. Biol., 195:247-259 (1987) provides a review of these. The preferred splice acceptor site is an efficient splice acceptor site which refers to a nucleic acid sequence encoding a splice

acceptor site wherein the efficiency of splicing of messenger RNA precursors having the splice acceptor site is between about 80 to 99% and preferably 90 to 95% as determined by quantitative PCR. The splice acceptor site may comprise a consensus sequence. In the mRNAs of higher eukaryotes, the 3' splice acceptor site occurs within the consensus sequence (U/C)₁₁NCAG:G. In the mRNAs of yeast, the 3' acceptor splice site is bounded by the consensus sequence (C/U)AG:. Padgett, *et al.*, *supra*.

As used herein "culturing for sufficient time to allow amplification to occur" refers to the act of physically culturing the eukaryotic host cells which have been transformed with the DNA construct in cell culture media containing the amplifying agent, until the copy number of the amplifiable gene (and preferably also the copy number of the product gene) in the host cells has increased relative to the transformed cells prior to this culturing.

The term "expression" as used herein refers to transcription or translation occurring within a host cell. The level of expression of a product gene in a host cell may be determined on the basis of either the amount of corresponding mRNA that is present in the cell or the amount of the protein encoded by the product gene that is produced by the cell. For example, mRNA transcribed from a product gene is desirably quantitated by northern hybridization or quantitative real-time PCR. Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, pp. 7.3-7.57 (Cold Spring Harbor Laboratory Press, 1989). Protein encoded by a product gene can be quantitated either by assaying for the biological activity of the protein or by employing assays that are independent of such activity, such as western blotting or radioimmunoassay using antibodies that are capable of reacting with the protein. Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, pp. 18.1-18.88 (Cold Spring Harbor Laboratory Press, 1989).

Modes for Carrying Out the Invention

Methods and compositions are provided for enhancing the stability and/or copy number of a transcribed sequence in order to allow for elevated levels of a RNA sequence of interest. In general, the methods of the present invention involve transfecting a eukaryotic host cell with an expression vector comprising both a product gene encoding a desired polypeptide and fused selectable genes.

Selectable genes and product genes may be obtained from genomic DNA, cDNA transcribed from cellular RNA, or by *in vitro* synthesis. For example, libraries are screened with

probes (such as antibodies or oligonucleotides of about 20-80 bases) designed to identify the selectable gene or the product gene (or the protein(s) encoded thereby). Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures as described in chapters 10-12 of Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the selectable gene or product gene is to use PCR methodology as described in section 14 of Sambrook *et al.*, *supra*.

A preferred method of practicing this invention is to use carefully selected oligonucleotide sequences to screen cDNA libraries from various tissues known to contain the selectable gene or product gene. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized.

The oligonucleotide generally is labeled such that it can be detected upon hybridization to DNA in the library being screened. The preferred method of labeling is to use ^{32}P -labeled ATP with polynucleotide kinase, as is well known in the art, to radiolabel the oligonucleotide. However, other methods may be used to label the oligonucleotide, including, but not limited to, biotinylation or enzyme labeling.

Sometimes, the DNA encoding the fused selectable genes and product gene is preceded by DNA encoding a signal sequence having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the expression vector, or it may be a part of the selectable gene or product gene that is inserted into the expression vector. If a heterologous signal sequence is used, it preferably is one that is recognized and processed (*i.e.*, cleaved by a signal peptidase) by the host cell. For yeast secretion the native signal sequence may be substituted by, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Pat. No. 5,010,182 issued 23 April 1991), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression the native signal sequence of the protein of interest is satisfactory, although other mammalian signal sequences may be suitable, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders, for example, the herpes simplex gD signal. The DNA for such precursor region is ligated in reading frame to the fused selectable genes or product gene.

As shown in Figure 1, the fused selectable genes are generally provided at the 5' end of the DNA construct and are followed by the product gene (which would be inserted into the linker site). Therefore, the full-length (non-spliced) message will contain, for example, the PURO-DHFR fusion as the first open reading frame and will therefore generate PURO-DHFR protein to allow selection of stable transfectants. The full length message is not expected to generate appreciable amounts of the protein of interest as the second AUG in a dicistronic message is an inefficient initiator of translation in mammalian cells (Kozak, J. Cell Biol., 115: 887-903 (1991)).

The fused selectable genes are positioned within an intron. Introns are noncoding nucleotide sequences, normally present within many eukaryotic genes, which are removed from newly transcribed mRNA precursors in a multiple-step process collectively referred to as splicing.

A single mechanism is thought to be responsible for the splicing of mRNA precursors in mammalian, plant, and yeast cells. In general, the process of splicing requires that the 5' and 3' ends of the intron be correctly cleaved and the resulting ends of the mRNA be accurately joined, such that a mature mRNA having the proper reading frame for protein synthesis is produced. Analysis of a variety of naturally occurring and synthetically constructed mutant genes has shown that nucleotide changes at many of the positions within the consensus sequences at the 5' and 3' splice sites have the effect of reducing or abolishing the synthesis of mature mRNA. Sharp, Science, 235:766 (1987); Padgett, *et al.*, Ann. Rev. Biochem., 55:1119 (1986); Green, Ann. Rev. Genet., 20:671 (1986). Mutational studies also have shown that RNA secondary structures involving splicing sites can affect the efficiency of splicing. Solnick, Cell, 43:667 (1985); Konarska, *et al.*, Cell, 42:165 (1985).

The length of the intron may also affect the efficiency of splicing. By making deletion mutations of different sizes within the large intron of the rabbit beta-globin gene, Wieringa, *et al.* determined that the minimum intron length necessary for correct splicing is about 69 nucleotides. Cell, 37:915 (1984). Similar studies of the intron of the adenovirus E1A region have shown that an intron length of about 78 nucleotides allows correct splicing to occur, but at reduced efficiency. Increasing the length of the intron to 91 nucleotides restores normal splicing efficiency, whereas truncating the intron to 63 nucleotides abolishes correct splicing. Ulfendahl, *et al.*, Nuc. Acids Res., 13:6299 (1985).

To be useful in the invention, the intron must have a length such that splicing of the intron from the mRNA is efficient. The preparation of introns of differing lengths is a routine matter,

involving methods well known in the art, such as *de novo* synthesis or *in vitro* deletion mutagenesis of an existing intron. Typically, the intron will have a length of at least about 150 nucleotides, since introns which are shorter than this tend to be spliced less efficiently. The upper limit for the length of the intron can be up to 30 kB or more. However, as a general proposition, the intron is generally less than about 10 kB in length.

The intron is modified to contain the fused selectable genes not normally present within the intron using any of the various known methods for modifying a nucleic acid *in vitro*. Typically, the fused selectable genes will be introduced into an intron by first cleaving the intron with a restriction endonuclease, and then covalently joining the resulting restriction fragments to the fused selectable genes in the correct orientation for host cell expression, for example by ligation with a DNA ligase enzyme.

The DNA construct is dicistronic, *i.e.* the fused selectable genes and product gene are both under the transcriptional control of a single transcriptional regulatory region. As mentioned above, the transcriptional regulatory region comprises a promoter. Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman *et al.*, J. Biol. Chem., 255:2073 (1980)) or other glycolytic enzymes (Hess *et al.*, J. Adv. Enzyme Reg., 7:149 (1968); and Holland, Biochemistry, 17:4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in Hitzeman *et al.*, EP 73,657A. Yeast enhancers also are advantageously used with yeast promoters.

Expression control sequences are known for eukaryotes. Virtually all eukaryotic genes have an AT-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is a CXCAAT region where X may be any nucleotide.

Product gene transcription from vectors in mammalian host cells is controlled by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, a retrovirus, hepatitis-B virus and most preferably Simian Virus 40 (SV40) or cytomegalovirus (CMV), from heterologous mammalian promoters, e.g. the actin promoter or an immunoglobulin promoter, from heat-shock promoters, and from the promoter normally associated with the product gene, provided such promoters are compatible with the host cell systems. Promoters endogenous to the host cell system, such as the CHO Elongation Factor 1 alpha promoter may also be used.

The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment that also contains the SV40 viral origin of replication. Fiers *et al.*, Nature, 273:113 (1978); Mulligan and Berg, Science, 209:1422-1427 (1980); Pavlakis *et al.*, Proc. Natl. Acad. Sci. USA, 78:7398-7402 (1981). The immediate early promoter of the human cytomegalovirus (CMV) is conveniently obtained as a *Hind*III E restriction fragment. Greenaway *et al.*, Gene, 18:355-360 (1982). A system for expressing DNA in mammalian hosts using the bovine papilloma virus as a vector is disclosed in U.S. 4,419,446. A modification of this system is described in U.S. 4,601,978. See also Gray *et al.*, Nature, 295:503-508 (1982) on expressing cDNA encoding immune interferon in monkey cells; , Reyes *et al.*, Nature, 297:598-601 (1982) on expression of human β -interferon cDNA in mouse cells under the control of a thymidine kinase promoter from herpes simplex virus, Canaani and Berg, Proc. Natl. Acad. Sci. USA, 79:5166-5170 (1982) on expression of the human interferon β 1 gene in cultured mouse and rabbit cells, and Gorman *et al.*, Proc. Natl. Acad. Sci. USA, 79:6777-6781 (1982) on expression of bacterial CAT sequences in CV-1 monkey kidney cells, chicken embryo fibroblasts, Chinese hamster ovary cells, HeLa cells, and mouse NIH-3T3 cells using the Rous sarcoma virus long terminal repeat as a promoter.

Preferably the transcriptional regulatory region in higher eukaryotes comprises an enhancer sequence. Enhancers are relatively orientation and position independent having been found 5' (Lainins *et al.*, Proc. Natl. Acad. Sci. USA, 78:993 (1981)) and 3' (Lusky *et al.*, Mol. Cell Bio., 3:1108 (1983)) to the transcription unit, within an intron (Banerji *et al.*, Cell, 33:729 (1983)) as well as within the coding sequence itself (Osborne *et al.*, Mol. Cell Bio., 4:1293 (1984)). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -

fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer (CMV), the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. See also Yaniv, Nature, 297:17-18 (1982) on enhancing elements for activation of eukaryotic promoters. The enhancer may be spliced into the vector at a position 5' or 3' to the product gene, but is preferably located at a site 5' from the promoter.

The DNA construct of the present invention has a transcriptional initiation site following the transcriptional regulatory region and a transcriptional termination region following the product gene (see, e.g., Figure 1). These sequences are provided in the DNA construct using techniques which are well known in the art.

The DNA construct normally forms part of an expression vector which may have other components such as an origin of replication (*i.e.*, a nucleic acid sequence that enables the vector to replicate in one or more selected host cells) and, if desired, one or more additional selectable gene(s). Construction of suitable vectors containing the desired coding and control sequences employs standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to generate the plasmids required.

Generally, in cloning vectors the origin of replication is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known. The 2 μ plasmid origin of replication is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (the SV40 origin may typically be used only because it contains the early promoter).

Most expression vectors are "shuttle" vectors, *i.e.*, they are capable of replication in at least one class of organisms but can be transfected into another organism for expression. For example, a vector is cloned in *E. coli* and then the same vector is transfected into yeast or mammalian cells for expression even though it is not capable of replicating independently of the host cell chromosome.

For analysis to confirm correct sequences in plasmids constructed, plasmids from the transformants are prepared, analyzed by restriction, and/or sequenced by the method of Messing *et*

al., Nucleic Acids Res., 9:309 (1981) or by the method of Maxam *et al.*, Methods in Enzymology, 65:499 (1980).

The expression vector having the DNA construct prepared as discussed above is transformed into a eukaryotic host cell. Suitable host cells for cloning or expressing the vectors herein are yeast or higher eukaryote cells.

Eukaryotic microbes such as filamentous fungi or yeast are suitable hosts for vectors containing the product gene. *Saccharomyces cerevisiae*, or common baker's yeast, is the most commonly used among lower eukaryotic host microorganisms. However, a number of other genera, species, and strains are commonly available and useful herein, such as *S. pombe* (Beach and Nurse, Nature, 290:140 (1981)), *Kluyveromyces lactis* (Louvencourt *et al.*, J. Bacteriol., 737 (1983)), *Kyarrowia* (EP 402,226), *Pichia pastoris* (EP 183,070), *Trichoderma reesia* (EP 244,234), *Neurospora crassa* (Case *et al.*, Proc. Natl. Acad. Sci. USA, 76:5259-5263 (1979)), and *Aspergillus* hosts such as *A. nidulans* (Ballance *et al.*, Biochem. Biophys. Res. Commun., 112:284-289 (1983); Tilburn *et al.*, Gene, 26:205-221 (1983); Yelton *et al.*, Proc. Natl. Acad. Sci. USA, 81:1470-1474 (1984)) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 (1985)).

Suitable host cells for the expression of the product gene are derived from multicellular organisms. Such host cells are capable of complex processing and glycosylation activities. In principle, any higher eukaryotic cell culture is workable, whether from vertebrate or invertebrate culture. Examples of invertebrate cells include plant and insect cells. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts such as *Spodoptera frugiperda* (caterpillar), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (fruitfly), and *Bombyx mori* host cells have been identified. See, e.g., Luckow *et al.*, Bio/Technology, 6:47-55 (1988); Miller *et al.*, in Genetic Engineering, Setlow, J.K. *et al.*, eds., Vol. 8 (Plenum Publishing, 1986), pp. 277-279; and Maeda *et al.*, Nature, 315:592-594 (1985). A variety of such viral strains are publicly available, e.g., the L-1 variant of *Autographa californica* NPV and the Bm-5 strain of *Bombyx mori* NPV, and such viruses may be used as the virus herein according to the present invention, particularly for transfection of *Spodoptera frugiperda* cells.

Plant cell cultures of cotton, corn, potato, soybean, petunia, tomato, and tobacco can be utilized as hosts. Typically, plant cells are transfected by incubation with certain strains of the bacterium *Agrobacterium tumefaciens*, which has been previously manipulated to contain the product gene. During incubation of the plant cell culture with *A. tumefaciens*, the product gene is

transferred to the plant cell host such that it is transfected, and will, under appropriate conditions, express the product gene. In addition, regulatory and signal sequences compatible with plant cells are available, such as the nopaline synthase promoter and polyadenylation signal sequences. Depicker *et al.*, J. Mol. Appl. Gen., 1:561 (1982). In addition, DNA segments isolated from the upstream region of the T-DNA 780 gene are capable of activating or increasing transcription levels of plant-expressible genes in recombinant DNA-containing plant tissue. EP 321,196 published 21 June 1989.

However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure in recent years (Tissue Culture, Academic Press, Kruse and Patterson, editors (1973)). Examples of useful mammalian host cell lines are monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham *et al.*, J. Gen Virol., 36:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); dp12.CHO cells (EP 307,247 published 15 March 1989); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Mather *et al.*, Annals N.Y. Acad. Sci., 383:44-68 (1982)); MRC 5 cells; FS4 cells; and a human hepatoma line (Hep G2).

Host cells are transformed with the above-described expression or cloning vectors of this invention and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw *et al.*, Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) may be used. General aspects of mammalian cell host system transformations have been described by Axel in U.S. 4,399,216 issued 16 August 1983. Transformations into yeast are typically carried out according to the method of Van Solingen *et al.*, J. Bact., 130:946 (1977) and Hsiao *et al.*, Proc. Natl. Acad. Sci. (USA), 76:3829

(1979). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used.

In preferred embodiments the DNA is introduced into the host cells using electroporation, lipofection or polyfection techniques. In a particularly preferred embodiment, the transfection is performed in a spinner vessel as illustrated by Example 3 or in some other form of suspension culture. Transfection performed in a spinner vessel is also referred to as "spinner transfection". Culturing the cells in suspension allows them to reach a cell density of at least about 5×10^5 /ml and more preferably at least about 1.5×10^6 /ml prior to transfection. See Andreason, J. Tiss. Cult. Meth., 15:56-62 (1993), for a review of electroporation techniques useful for practicing the claimed invention. It was discovered that these techniques for introducing the DNA construct into the host cells are preferable over calcium phosphate precipitation techniques insofar as the latter could cause the DNA to break up and form concatemers.

The mammalian host cells used to express the product gene herein may be cultured in a variety of media as discussed in the definitions section above. The media is formulated to provide selective nutrient conditions or a selection agent to select transformed host cells which have taken up the DNA construct (either as an intra- or extra-chromosomal element). To achieve selection of the transformed eukaryotic cells, the host cells may be grown in cell culture plates and individual colonies expressing one or both of the selectable genes (and thus the product gene) can be isolated and grown in growth medium under defined conditions. The host cells are then analyzed for transcription and/or transformation as discussed below. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)), dot blotting (DNA or mRNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Various labels may be employed, most commonly radioisotopes, particularly ^{32}P . However, other techniques may also be employed, such as using biotin-modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorescence, enzymes, or the like. Alternatively, antibodies may be employed that can recognize specific

duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. With immunohistochemical staining techniques, a cell sample is prepared, typically by dehydration and fixation, followed by reaction with labeled antibodies specific for the gene product coupled, where the labels are usually visually detectable, such as enzymatic labels, fluorescent labels, luminescent labels, and the like. A particularly sensitive staining technique suitable for use in the present invention is described by Hsu *et al.*, Am. J. Clin. Path., 75:734-738 (1980).

In the preferred embodiment protein expression is measured using ELISA as described in Example 1 herein.

The product of interest preferably is recovered from the culture medium as a secreted polypeptide, although it also may be recovered from host cell lysates when directly expressed without a secretory signal. When the product gene is expressed in a recombinant cell other than one of human origin, the product of interest is completely free of proteins or polypeptides of human origin. However, it is necessary to purify the product of interest from recombinant cell proteins or polypeptides to obtain preparations that are substantially homogeneous as to the product of interest. As a first step, the culture medium or lysate is centrifuged to remove particulate cell debris. The product of interest thereafter is purified from contaminant soluble proteins and polypeptides, for example, by fractionation on immunoaffinity or ion-exchange columns; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel electrophoresis using, for example, Sephadex G-75; chromatography on plasminogen columns to bind the product of interest and protein A Sepharose columns to remove contaminants such as IgG.

The following examples are offered by way of illustration only and are not intended to limit the invention in any manner. All patent and literature references cited herein are expressly incorporated by reference.

EXAMPLE 1

2C4 production using the fusion construct expression vector

Vectors related to those described by Lucas et al (Lucas BK, Giere LM, DeMarco RA, Shen A, Chisholm V and Crowley C. High-level production of recombinant proteins in CHO cells using a dicistronic DHFR intron expression vector. (1996) *Nucleic Acids Res.* 24(9), 1774-1779.), which contain an intron between the SV40 promoter and enhancer and the cDNA that encodes the polypeptide of interest, were constructed. The intron is bordered on its 3' and 5' ends, respectively, by a splice donor site derived from cytomegalovirus immediate early gene (CMVIE), and a splice acceptor site from an IgG heavy chain variable region (V_H) gene (Eaton *et al.*, *Biochem.*, 25:8343 (1986)). The splice sites selected provide slightly inefficient splicing such that only about 90% of the transcripts produced are intron free. Previous studies have demonstrated that when a selectable marker such as DHFR is integrated within this intron, as in the plasmid pSV.ID, marker gene transcription proceeds from any unspliced transcripts, providing a highly efficient means of maintaining linkage between the expression of the marker gene and the cDNA of interest as well as enhanced product expression relative to expression of the marker gene.

Vectors containing a murine puromycin/DHFR fusion sequence in the intron following the SV40 promoter elements were constructed by linearizing a pSV.IPUR plasmid, which contained the puromycin resistance gene in an intron following the SV40 promoter/enhancer (pSV.IPUR, Figures 1 and 2), with Hpa I immediately following the end of the puromycin ORF. A 564 bp PCR fragment containing the entire coding region for the murine DHFR gene was subsequently ligated into this linearized vector 3' of the puromycin resistance gene. The stop codon TAG between the puromycin resistance gene and the DHFR gene was deleted by site-directed mutagenesis resulting in a pSV.I plasmid containing a Puro/DHFR fusion gene within the intron of the expression cassette (pSV.IPD, Figures 1 and 4).

The cDNA of the Heavy chain (HC) and light chain (LC) sequences of an anti-HER2 Neu antibody, 2C4, were inserted into pSV.IPD as shown in Figure 6. The sequence of the resulting pSV.IPD.2C4 vector is shown in Figure 7. Data collected using the pSV.IPD.2C4 vector are shown in Table 2.

Additionally, a vector containing only a murine DHFR sequence within the intron (pSV.ID) was prepared. The DNA sequence for the pSV.ID vector is shown in Figure 3. The preparation of such vectors is disclosed in U.S. Patent No. 5,561,053, which is herein incorporated by reference. Into that vector, the HC and LC sequences of monoclonal antibodies to VEGF were inserted. The sequence of the resulting pSV.ID.VEGF vector is shown in Figure 5.

Plasmid DNA's that contained either the Puro/DHFR fusion sequences in the intron or murine DHFR alone preceding cDNA sequences for HC and LC of 2C4 and anti-VEGF, respectively were introduced into CHO DHFR minus cells by lipofection. Briefly, for transfection, 4 million CHO DUX-B11 (DHFR minus) were seeded in 10 cm plates the day before transfection. On the day of transfection, 4 ug DNA was mixed with 300 ul of serum free medium and 25 ul of polyfect from Qiagen. The mixture was incubated at room temperature for 5 to 10 minutes and added to the cells. Cells were fed with fresh glycine, hypoxanthine and thymidine-free (GHT-free) medium and twenty-four hours later, were trypsinized and selected in fresh GHT-free medium with 0 – 5 nM of methotrexate (MTX) in order to select for stable DHFR+ clones. Approximately 300 – 400 individual clones were selected in this first round of screening for measurement of protein expression levels. Clones from each vector which expressed the highest levels of antibody were then re-exposed to higher levels of methotrexate to affect a second round of gene amplification and selection. The screening process was repeated on all available clones, the highest of which were exposed to a third round of amplification. The methotrexate concentrations used during amplification using the pSV.ID-derived vector was 50 to 1000 nM in the 2nd round and 200 to 1000 nM in the 3rd round. These concentrations are typically required to achieve growth-limiting toxicity, which is required to achieve sufficient selective pressure for gene amplification. Concentrations required to reach this same degree of toxicity using the pSV.ID-derived vectors were remarkably lower.

The level of antibody expression was determined by seeding cells in 1 ml of serum-free F12:DMEM-based media supplemented with protein hydrolysate and amino acids in 24 well dishes at 3×10^6 cells/ml or in 100 ul of similar media in individual wells of a 96 well plate. Growth media was collected after 3-4 days and titers were assayed by an ELISA directed towards the intact IgG molecule. In experiments where cells were not seeded at equal cell densities, a fluorescent measure of viable cell number was performed on each well in order to normalize expression data. An Intact IgG ELISA was performed on microtiter plates which used a capture

antisera directed to framework Fab residues common in both antibodies. Media samples were added to the wells followed by washing and a horseradish peroxidase labeled second antibody directed towards common framework Fc residues was used for detection.

Table 2 presents expression level distributions of clones isolated during each round of screening of anti VEGF clones, which resulted from transfection with the plasmid containing only the DHFR sequence in the intron (pSV.ID.aVEGF), and 2C4 clones that were created using the Puro/DHFR fusion sequence in the same intron (pSV.IPD.2C4). The distribution of expression levels seen in the case of anti VEGF is typical of the performance of the vector containing only the murine DHFR gene in the intron (pSV.ID). All isolates identified in the first and second rounds of screening have relatively low expression levels. In the initial selection round, no clones with expression above 5 were isolated. At least three rounds of amplification are required to identify clones capable of specific productivity greater than 50. The 2C4 clones were screened after the first exposure to methotrexate (0-2.5 nM) and the most productive of these were exposed to a second round of amplification in 10-25 nM MTX. Cells surviving this amplification were pooled and exposed to 3rd round amplification prior to selection for further screening. In contrast to the pSV.ID vector, using the pSV.IPD vector, clones with an expression level of up to 25 were identified even in the first round of screening. Clones with an expression level greater than 25 represented 95% of the population after their third round of amplification and screening.

The data from Example 1 indicates that use of the Puro/DHFR fusion protein as the selectable marker allows for faster, more efficient isolation of highly productive CHO clones using significantly lower levels of methotrexate. The data suggests that exposure to low concentrations and stepwise increments in methotrexate allow for the efficient initial selection of highly expressing clones and subsequent gene amplification. Exposure to excessively high concentrations of methotrexate or large incremental increases in exposure often does not yield increases in gene expression since cells rapidly acquire methotrexate resistance through non-gene amplification mechanisms. Importantly, the data also shows that the Puro/DHFR fusion protein provides an unexpectedly impaired activity of the DHFR gene product or an enhanced sensitivity to methotrexate, which results in a highly stringent initial selection step, and allows efficient gene amplification at concentrations of methotrexate not frequently associated with the acquisition of drug resistance through alternative mechanisms. The ability to select cells which have incorporated the plasmid either in the presence of puromycin or methotrexate, prior to initiating exposure to

methotrexate also provides a means of transferring this efficient system to DHFR (positive) host cells.

For Example 1 the structure of the expressed antibody has been extensively characterized. The proteins generated from the pSV.IPD are indistinguishable from the antibody produced by the pSV.ID vector, with no apparent increase of free heavy or light chain expressed by the pool.

TABLE 2. PERCENTAGES OF pSV.IPD.2C4 CLONES ISOLATED AT VARIOUS EXPRESSION LEVELS AFTER MTX EXPOSURE¹

Expression Level ²	pSV.ID.aVEGF 1st Rd	pSV.IPD.2C4 1st Rd	pSV.ID.aVEGF 3rd Rd	pSV.IPD.2C4 3rd Rd
<1	71	16	0	0
1-5	29	67	0	0
5-10	0	14	2	3
10-25	0	3	15	4
25-50	0	0	35	21
50-100	0	0	46	61
100-150	0	0	2	3

¹ MTX concentration for Control SD vector = 0-10 nM 1st round, 50-1000 nM 2nd round, 200-1000 nM, 3rd round. SD- Puro/DHFR vector = 2.5 nM 1st round, 25 nM 2nd round, 100 nM 3rd round.

² Expression levels are in mg/ml or (mg/ml)/Fluorescent Unit

This example demonstrate the general applicability of the Puro/DHFR fusion sequence for selection of highly productive recombinant cell lines following minimal exposure to MTX.

EXAMPLE 2

Recombinant protein production using a pSV.I construct containing DHFR and a fusion gene other than Puro

Constructs can also be produced that contain a fusion sequence of an alternative selectable marker and DHFR within an intron region as described in Example 1. For instance

starting with the vector pSVID, the coding sequences for the neomycin resistance gene (Neo), hygromycin resistance gene (Hygro), glutamine synthase (GS), thymidine kinase (TK) or zeocin (Zeo) could be inserted in frame with the start site of the murine DHFR sequence contained within the intron. The stop codon of this inserted gene would then be removed using site directed mutagenesis according to example 1. Depending upon the phenotype of the host cell selected, cells incorporating the plasmid could then be selected using either GHT-free or MTX containing media as described in examples 1-3 or using an appropriate quantity of the alternative selective agent. Gene expression by the resulting clones could then be amplified in the presence of increased levels of methotrexate.

EXAMPLE 3

Direct Selection with plasmids SV.IPD.HP and CMV.IPD.HP after spinner transfection

DP12 CHO cells were grown in growth medium with 5% FBS (fetal bovine serum) and 1X GHT (glycine, hypoxanthine and thymidine). The process typically took about 4 days. On day 1, cells were seeded at 4×10^5 /ml in 400 ml growth medium in a 500 ml spinner vessel and grown for 2 days at 37 °C. On day 3, the exponentially grown cells were seeded at 1.5×10^6 cells/ml in a 250 ml spinner vessel containing 200 ml of growth medium plus 5% FBS and 1X GHT. The cells were grown for 1 to 2 hours at 37 °C before transfection. During that time, serum-free growth medium and 1X GHT was warmed to 37 °C. 400 µg plasmid construct DNA and 1 ml of Lipofectamine 2000® (Qiagen) were separately diluted into 25 ml of warm serum-free medium in 50 ml Falcon tubes. The solutions in the tubes were combined and incubated at room temperature for 30 minutes. The cells were then transfected with plasmid constructs pSV.IPD.HP and pCMV.IPD.HP, which constructs are illustrated in Figures 13 and 14, respectively. At the end of incubation, the cells were transfected by adding all 50 ml of the mixture of diluted plasmid construct and Lipofectamine 2000® to the 250 ml spinner vessel containing cells in serum-free medium, and the cells continued to grow at 37 °C for about 24 hours. On day 4, 250 ml of transfected cells were centrifuged at 1000 rpm for 5 minutes to collect the pellet. The transfection efficiency was monitored by transfecting cells with a GFP plasmid followed by FACS analysis 24 hours after transfection. The transfection efficiency with this protocol was typically approximately 55 to 70 % in CHO cells as shown in Figure 8.

After the transfection, cells were centrifuged to collect the pellet. The pellet was then resuspended in growth medium containing methotrexate (MTX) ranging from 10 to 100 nM for either SV40 or CMV based constructs. Approximately 100 clones survived the direct selection. Cell growth medium was changed every 3 to 4 days. At approximately 2 weeks after transfection, individual clones were picked and grown in 96-well plates in growth medium containing MTX. Heterologous polypeptide expression levels were evaluated by ELISA. Figures 10-1, 10-2, and 11 show the results from 25 nM and 50 nM MTX selection. Figure 9 shows heterologous polypeptide expression levels of clones from a traditional 10 nM MTX selection where the cells were not transfected in a spinner flask.

It took about 1 week for cells to grow confluent in a 96-well plate. When they were confluent, the growth medium was removed and commercially available enriched cell culture medium (which includes 1x GHT but no MTX) was added into each well. On the day after adding the commercially available enriched cell culture medium, the plate was incubated at 33 °C for 5-6 days before performing an ELISA assay to quantitate the amount of humanized monoclonal antibody produced by the cells. ELISA was typically performed with serial dilutions of the commercially available enriched cell culture medium. Results from a humanized monoclonal antibody production were shown in Figures 9, 10-1, 10-2 and 11.

The four clones producing the greatest amount over 100 µg/ml of intact IgG based on direct selection at 25 nM MTX using a CMV-based construct were scaled up from a 96-well plate to a 6-well plate and then to a 10 cm plate. Cells were seeded at 3×10^5 /ml in 200 ml volume in a 250 ml spinner vessel in serum-free growth medium with 2 µg/ml human insulin and 1X Trace Elements (TE). Cells were initially passaged at either two- or three-day intervals with medium exchange. Then they were passaged at either three- or four-day intervals for about 6 weeks before bioreactor evaluation. At each passage time, cell viability and count number were monitored. To determine the cell growth after serum-free adaptation, a spinner vessel growth experiment was performed. Cells were seeded at 3×10^5 cells/ml into 400 ml of growth medium with 2 µg/ml recombinant human insulin and 1X TE in a 500 ml spinner vessel on day 1. On each day, packed cell volume (PCV) was monitored until day 5. PCVs reached between 0.4 % to 0.6% by day 4. Two serum-free adapted clones from 25 nM MTX selection with CMV-based construct were evaluated in bioreactors. Two liter bioreactors with commercially available

enriched cell culture medium were run for a total of 14 days. The data from the titer evaluation is shown in Figure 12.

An ELISA assay of clones surviving the direct selection shows that the best clones coming out of the method described in this example produce as much product of interest as highly amplified clones from a traditional method. See Figure 16. Evaluations of 2 clones from the direct selection shows that those clones produce about 1g/L of a product of interest in a bioreactor process. Since those clones were generated from one step of a direct selection immediately after transfection, it only takes about 5 to 6 weeks to generate a stable cell line producing 1g/L of a product of interest in a bioreactor leading to significant timeline reduction, about 3 months, which is critical for efficiency of product development.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the examples presented herein, since the exemplified embodiments are intended as illustrations of certain aspects of the invention and any functionally equivalent embodiments are within the scope of this invention. The examples presented herein are not intended as limiting the scope of the claims to the specific illustrations. Indeed, various modifications of the invention, in addition to those shown and described herein and which fall within the scope of the appended claims, may become apparent to those skilled in the art from the foregoing description.

CLAIMS

What is claimed is:

1. A method of producing a host cell capable of producing a product of interest, comprising:

transfecting a host cell culture with a DNA construct comprising a transcriptional regulatory region, a fused selectable gene sequence and a gene encoding a product of interest;

directly culturing the transfected host cells in a selective medium;

allowing the host cells to grow in the selective medium for a sufficient time to allow amplification of gene encoding the product of interest to occur; and

selecting a host cell clone that is capable of producing at least about 250mg/l of the product of interest.
2. A method of claim 1 wherein the selective medium contains at least about 25nM methotrexate.
3. A method of claim 1 wherein the selective medium contains at least about 50nM methotrexate.
4. A method of claim 1 wherein the host cell is a CHO cell.
5. A method of claim 1 wherein the product of interest is a protein selected from the group consisting of an antibody, enzyme, hormone, lipoprotein, clotting factor, anti-clotting factor, cytokine, viral antigen, chimeric protein, transport protein, regulatory protein, homing receptor, and addressin; or a fragment of said protein.
6. A method of claim 1 wherein said product of interest is a humanized antibody.
7. A host cell produced according to the method of claim 1.

8. A method of producing a product of interest, comprising culturing a host cell produced according to the method of claim 1 under conditions suitable to cause expression of the product of interest in an amount at least about 250mg/l.

9. A method of claim 1 wherein the DNA construct comprises, in order 5' to 3':

a) a transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene;

b) a transcriptional initiation site;

c) a fused selectable gene sequence positioned within an intron defined by a 5' splice donor site comprising a splice donor sequence such that the efficiency of splicing messenger RNA having said splice donor sequence is between about 80% and 99% as determined by PCR, and a 3' splice acceptor site;

d) a product gene encoding a product of interest; and

e) a transcriptional termination site.

10. The method of claim 9 further comprising recovering the product of interest from the culture.

11. A method of claim 9 wherein the transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene is driven by a SV40 promoter.

12. A method of claim 9 wherein the transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene is driven by a CMV promoter.

13. A cell culture composition comprising a host cell according to claim 9 and at least about 250mg/l of the product of interest.

14. A method of producing a host cell capable of producing at least about 250mg/ml of a product of interest comprising transfecting a host cell with a DNA construct comprising in order from 5' to 3':

a) a transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene;

b) a transcriptional initiation site;

c) a fused selectable gene sequence positioned within an intron defined by a 5' splice donor site comprising a splice donor sequence such that the efficiency of splicing messenger RNA having said splice donor sequence is between about 80% and 99% as determined by PCR, and a 3' splice acceptor site;

d) a product gene encoding a product of interest; and

e) a transcriptional termination site;

wherein the transfection is performed in suspension culture.

15. A method of claim 14, wherein the DNA construct is introduced into the host cells by lipofection.

16. A method of claim 14 wherein said transfection is performed in a spinner vessel.

17. The method of claim 14 wherein the suspension culture has cell density of at least about 5×10^5 /ml at the time of transfection.

18. The method of claim 14 wherein the suspension culture has a cell density of at least about 1.5×10^5 /ml at the time of transfection

19. A method of claim 15 wherein the product of interest is selected from the group consisting of an antibody, enzyme, hormone, lipoprotein, clotting factor, anti-clotting factor, cytokine, viral antigen, chimeric protein, transport protein, regulatory protein, homing receptor, and addressin and a fragment of any of said product of interest.

20. A method of rapidly selecting a host cell producing a product of interest, comprising:

transfecting a host cell culture with a DNA construct comprising a transcriptional regulatory region, a fused selectable gene sequence and a gene encoding a product of interest;

directly culturing the transfected host cells in a selective medium; and

allowing the host cells to grow in the selective medium for a sufficient time to allow amplification of gene encoding the product of interest to occur.

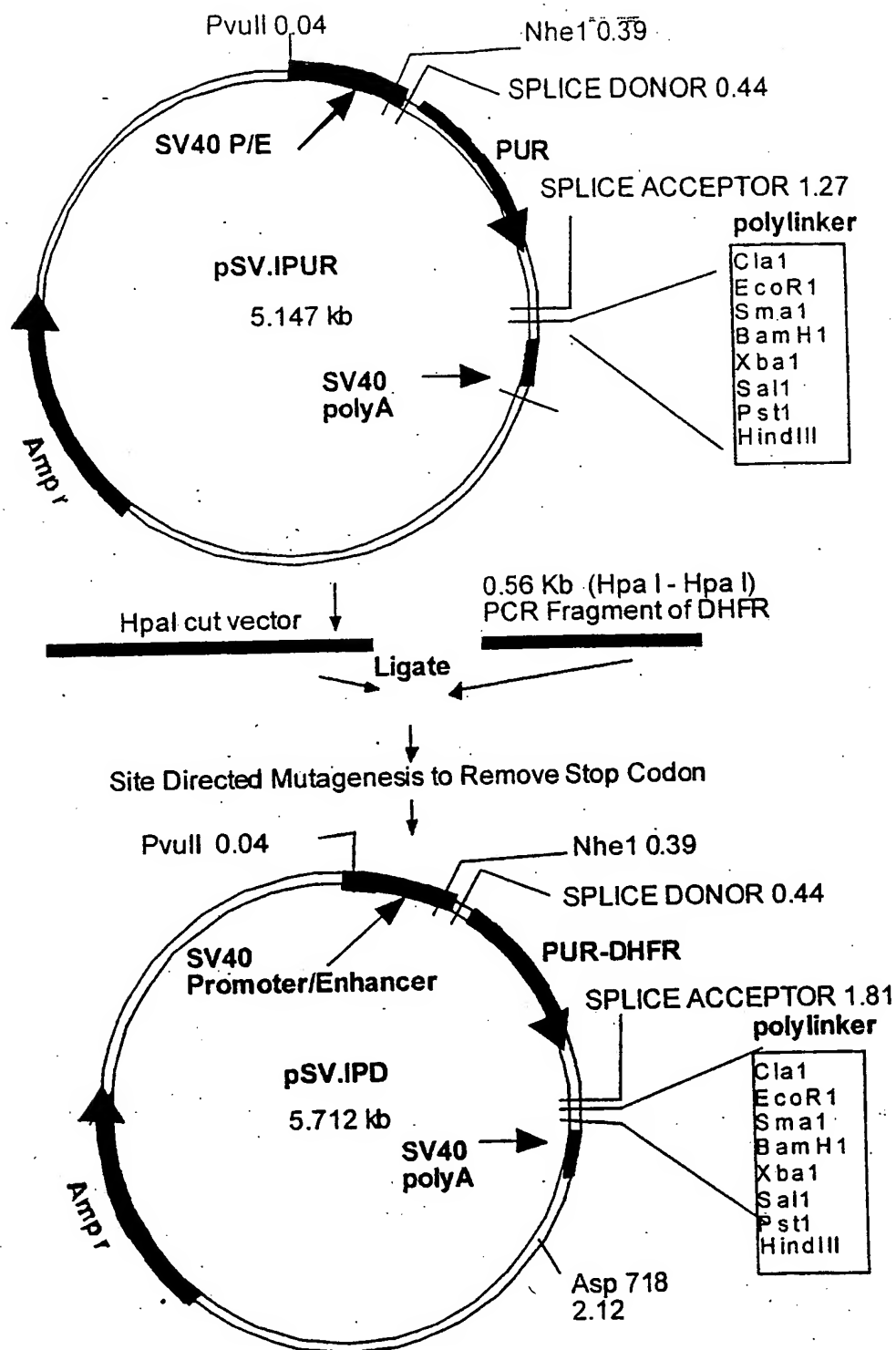


Figure 1. Construction of pSV.IPD Plasmid

Figure 2
psv.IPUR
length: 5147 (circular)

```

1  TTCGAGCTCG CCGACATTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT CAGTTAGGTT GTGGAAGTCC CCAAGCAAGA CCAUCAAUA/A
   AAGCTCGAGC GGGCTGTAACT TAATAACTGA TCTCAGCTAG CTGTGCACAC CTTACACACA GTCAATCCCA CACCTTTTCAG GGGTCCAGG GGTTCCTCCT
101 GAAGTATGCA AAGCATGCTAT CTCATTAGT CAGCAACCGG GTGTGGAAG TCCCCAGGCT CCCACGAGG CAGAAGTATG CAAAGCATGC ATCTCAATTA
   CTTTCATACGT TTCGTACGTA GAGTTAATCA GTCGTTGGTC CACACCTTTC AGGGTCCGA GGGGTCTCTC GTCTTCATAC GTTTCGTAGC TAGAGTTAAT
201 GTCAGCAACC ATAGTCCCGC CCCTRACTCC GCCATCCCG CCCTRACTC CGCCCACTTC CGCCCATTTCT CGGCCCATG GCTGALTAAT TTTTATTATT
   CAGTCGTGG TATCAGGGCG GGGATTGAGG CGGGTAGGC GGGGTTGAG CGGGTCAAG CGGGGTAC GAGCTGATTA AAAAAATTA
301 TATGCGAGG CCGAGCGCGC CTGGGCTCT GAGCTATTCC AGAAGTAGTG AGGAGGCTTT TTTGGAGGCC TAGCTTTTTC CAAAAAGCTTA GCTTATTCCTC
   ATACGCTCTC GGTCCGGCG GAGCCGAGG CTCGATAAGG TCTTCATCAC TCCTCCGAA AARCTCCGG ATCCGAAAC GTTTTTCGAT CGAATAGGCG
401 CCGGGAACGG TGCATTGGAA CGCGATTCC CGTGCCAAAG AGTGACATAA GTACCGGCTA TAGAGGACT AGTCCACCAT GACCGAGTAC AAGCCACACG
   GGGCTTGGC ACCTAACCTT GCGCTTAAGG GGCACGGTTC TCACTGCATT CATGGCGCAT ATCTCGGTGA TCATGTGTA CTGGCTCATG TTTTGGTATC
501 TCGGCTCTCG CACCCGCGAC GAGCTCCCGC GGGCGGTACG CACCTCGCC GCGCGTTCG CCGACTACCC GCGCACGCGC CACACCTTCC ACCUUAATC
   ACGCGAGCG GTGGGCGCTG CTCGAGGGG CCGGGCATGC GTGGAGCGG CGGGCAAGC GGTGATGG GCGGTGCGCG GTGTGGCAGC TGGGCTTGGC
601 CCACATCGAG CGGCTCACCG AGCTGCAAGA ACTCTTCTCT ACGCGGTGCG GGCTCGACAT CGGCAAGGTG TGGGTGCGCG ACGAGGCGC CAGTGTGCGC
   GGTGTAGCTC GCGCAGTGGC TCACAGTCTT TGAGAGGAG TGCGCGGAGC CCGAGCTGTA GCGCTTCCAC ACCCAGCGCC TGCTGCCCGG GCGCCACCG
701 GTCTGACCA CCGCGGAGG CGTCGAGCG GGGCGGTGT TCGCCGAGAT CGGCCCGCG ATGCCGAGT TGAGCGGTT CCGGTGCCC GCGCAGTAAC
   CAGACCTGT GCGGCTCTC GCAGCTTCGC CCGCGCACCA AGCGGCTCTA CCGCGGCGG TACCGGCTCA ACTCCCAAG GCGCGACCG CCGCTTCTTC
801 AGATGGAAG CTTCTTGGG CCGCACCGG CCAAGGAGC CGGTGCTTC GTGSCCACC GTGCGCTCTC GCGCGACCG CAGGCAAGG GTCGCGTCC CAGTACCTTC
   TCTACCTTCC GGAGGACCG GCGGTGCGG GGTTCCTCG GCGACCAAG GACCGGTGGC AGCGCGAG GCGGCTGGT GTCGCGTTC CAGTACCTTC
901 CGCGGTCTG CTCCTCGGAG TGGAGGGCG CAGCGCGCG GGGGTGCGC CTTTCTTGA GACTTCCCG CCCCGCAACC TCCCTTCTTA CCAAGCGCTC
   GCGGACGAC GAGGGGCTC ACCTCGCGG GCTCGCGGG CCCACGCGG GGAAGGACT CTGGAGGCG GGGGCTTGG AGGCAAGAT GCTCCCGGAG
1001 GCGTTACCG TCACCGCGG CGTCGAGTGC CCGAAGGAG GCGGACTCT GTGCATGACC CGCAAGCGCG GTGCTGAGT TAACCTGCTC CTTTCTTAAAG
   CCGAAGTGC AGTGGCGGT GCAGCTCAG GGTTCCTCG CCGCTGAGC CACCTACTGG CCGTTCGGG CACGACTCA ATTGACGAG GAGGATTTTC
1101 CTATGATTTT TTATAGACC ATGGACTTT TGCTGGCTTT AGATCCCTT AGATCGGTA GAGCGAGT ACAATTAATA CATAACCTTA TCTATCATAC
   GATACCTAAA ATATTCTGG TACCTCGAA ACAGCGAAA TCTAGGGAA CCGAAGCAAT CTGCGTCTGA TGTAAATAT GTATTGGGAT ACATATCATC
1201 ACATACGATT TAGGTGACAC TATAGATAAC ATCCACTTTG CTTTCTCTC CACAGGTGTC CACTCCCGG TCCCACTGCA CTTCTCTTCT ATCTATTAATA
   TGTATGCTAA ATCCACTGTG ATATCTATTG TAGTGAAC GGAAGAGAG GTGTCCACAG GTGAGGGTCC AGGTTGAGT GGAGCCAAGA TAUCTAAATT
1301 TTCCCGGGG ATCTCTAGA GTGACCTGC AGAAGCTTCG ATGCGCGCA TGGCCCACT TGTATTATTC ACCTTATTAAT GCTTATTAAT AAGGATATTC
   AAGGGGCCC TAGGATCTC CAGCTGGAGC TCTTCGAGC TCTTCGAGC TACCGGCGT ACCGGGTGA ACAATAACG TCGAATATTA CCAATGTTTA TTTCTTATTC

```

Figure 2-1

1401 CATCAAAAT TTCACAAATA AAGCATTTT TTACATGCAAT TCTAGTGTG GTTGTGCCAA ACTCATCAAT GTATCTTATC ATGTCGTGAT CUATCGGHA
 CTAGTGTGTTA AAGTGTGTTAT TTCTGTAATA AAGTGACGTA AGATCAACAC CAACACAGGT TGAGTAGTTA CATAGATAG TACAGACCTA CTATGCCCC

 1501 TTAATTGGC GCAGCACCAAT GGCCTGAAAT AACCTCTGAA AGAGGAACCTT GGTAGTAGAC CTCTGTGAGG GGAAGAACC AGCTGTGAA TGTGTGTAT
 AATTAAGCG CGTGTGGTA CCGGACTTTA TTGGACACTT TCTCTGTGAA CCAATCCATG GAAGACTCCG CCTTCTTGG TCGACACCTT ACACACACTC

 1601 TTAGGGTGT GAAAGTCCCG AGGCTCCCCA GCAGGCAGAA GTATGCAAG CATGATCTC AATAGTCAG CAACACAGTG TGAAGATCC CAGGCTCCC
 AATCCACAC CTTCAGGGG TCCGAGGGT CCGCTCTCTT CATACGTTT GTACGTAGG TTAATCAGTC GTTGTCCAC ACCTTTCAGG GGTCCGAGG

 1701 CAGCAGGAG AAGTATGCA AGCATGCAT TCAATTAGTC AGCAACATA GTCCCGGCC TAACTCCGC CATCCCGCC TAACTCCGC CCACTTCCGC
 GTCGTCCGT TCCATACGTT TCGTAGTAG AGTTATACAG TCGTTGGTAT CAGGGCGGG ATTGAGCGG GATTGAGCG GGTCAAGGCG

 1801 CCAATCTCG CCCATGGCT GACTAATTTT TTTTATTTAT GCAGAGCGG AGGCGGCTC GGCTCTGAG CTATTCAGA ACTAGTGGG AUGCTTTT
 GGTAGAGCG GGGTACCGA CTGATTAAAT AATTAATAA CCGTCCGCG TCCGCGGAG CCGAGACTC GATPAGGTCT TCATCACTCC TCCGAAAAA

 1901 GGAGGCTAG GCTTTTGCA AAGCTGTTA CCGTAGCGG CGCTTAAT AAGCGCGCC ATTTAATCC TGCAGTAAAC AGCTTGGAC TGGCGTCTG
 CCGCGGATC CGAATACGTT TTTCGACAT GGAGCTCGC GCGGATTA TTCCGCGCG TAAATTTAG ACCTCCATTG TCGAACCGTG ACCGCGCA

 2001 TTACAACGT CGTAGCTGG AAGCCCTGG CGTTACCCAA CTTAATCGCC TTGCAGACA TCCCGCTTC GCAGCTGGC GTATAGCGA AGAGGCGGCG
 AATGTTGCA GCATGACCC TTTCGGACC GCAATGGGT GAATAGCGG AAGCTGTGT AGGGGGAAG CGGTGACCG CATATCGCT TCTCGGCGG

 2101 ACCGATCGC CTTCACAA CTTGCTAGC CTGATGGG AATGGCGCT GATGCGTAT TTTCTCTTA CGCATCTGT CGGTATTTCA CACUUCATAC
 TGGCTAGCG GAAGGTTGT CAACGATCG GACTTACCG GATCGGCGA CTACGCAATA AAGAGGAAT GCGTAGACAC GGCATAAAGI GTGGGTATG

 2201 GTCAAGCA CCAATAGTAG CCGCTCTAG CCGCGATTA AGCGCGCGG GTGTGTGTGT TAGCGGAGC GTACCGCTA CACTTGGCAG CUCCTTAGG
 CAGTTTCGT GGTATCATG GCGGACATC GCGCGTAT TCGCGCGCC CACACACCA ATGCGGTG CACTGGCGT GTGAACGCTC GCGGATCGC

 2301 CCGCTCTCT TCGCTTCTT CCGTCTCTT CTCGCGAGT TCGCGCGCTT TCGCGCGCTT TCGCGCGCTT TCGCGCGCTT TCGCGCGCTT TCGCGCGCTT
 GCGCGAGGA AGCGAAGAA GCGAGGTA GCGCGTGA AGCGCGGTA AGCGCGGTA AGCGCGGTA AGCGCGGTA AGCGCGGTA AGCGCGGTA

 2401 CTTTACGCA CCGTACCGC AAAAACTTG ATTTGGTGA TGGTTCAGT AGTGGCGCT CCGCTGATA GCGGTTTTT GCGGTTTTT GCGGTTTTT
 GATATGCGT GGAGCTGGG TTTTGTGAC TAAACCACT ACCAAGTGA TCACCGGTA TCACCGGTA TCACCGGTA TCACCGGTA TCACCGGTA TCACCGGTA

 2501 CAGCTCTTT AATAGTGGAC TCTGTGTTCA AACTGCAAC AACTCAACC CTATCTCGG CTATCTCTTT GATTTATAAG GATTTTGGC GATTTTGGC
 GTGAGAAA TTATCACCTG AGACAGGT TTGACCTGT TGTGAGTTG GATAGAGCC GATAGAAA CTAAATATTC CCTAARACG CTAAAGCGG

 2601 TATTGTTAA AAAATGAGT GATTTAACA AATTTAAG CGAATTTTA CAAATATTA AGTTTACAA TTTTATGGT CACTCTCGT ACATCTCTT
 ATACCAAT TTTTACTGA CTAAATGTT TTTAATGCG GGTAAATTT GTTTTATAAT TGAATATTT TGAATATTT TGAATATTT TGAATATTT

 2701 CTGATCGCC ATAGTTAAG CAATCCGCT ATCGCTACG GACTGGTCA TGGCTCGCC CCGACACCG CCAACACCG CTGACCGCC CTGACCGCC
 GACTACGCG TATCAATCG GTTGAGGCA TAGCGATGA CTGACCACT ACCGAGCGG GGTGTGGC GGTGTGGC GGTGTGGC GGTGTGGC GGTGTGGC

 2801 TGTCTGCTC CGGATCCCG TTACAGCAA GCTGTACCG TCTCGGGAG CTGATGCTG CAGAGTTTT CAGAGTTTT CAGAGTTTT CAGAGTTTT CAGAGTTTT
 ACAGACGAG GCGTAGGCG AATGCTGTT CACACACCG AGAGGCGCT CAGCTACCA GTCTCCAAA GTGGGAGTAG TGGTTTTGCG CUCCTCTCTA

 2901 ATTCTGTAG ACMAAGGCG CTCGTGATAC GCTATTTT ATAGTTAAT GTCATGATA TAATGTTT TTAGAGCTCA GCTGAGCTT TTTTAAAAA
 TAAGACTTC TGCTTTCCG GAGCACTAG CCGATAAAA TATCCAAAT CAGTACTAT ATTACCAAG AATCTCCAGT CCACCTCTAA AUGCTTTT

 3001 TGTGCGGCA ACCCTATTT GTTATTTT CTAAATACAT TCAATATGT ATCGCTCAT GAGACAATA CCGTCAATA TCGTCAATA ATATCAATA

Figure 2-2

ACACCGCGCT TGGGATATAA CAAATATAAA GATTATATGA AGTTTATACA TAGGCGAGTA CTCTGTATTT GGGACTATTT ACAGAGTTAT TATTAACATTTT
3101 AGGAAGAGTA TGAGTATTCA AGATTTCGGT GTGCGCCTTA TTCCCTTTT TGGGCAATTT TGCGGCATTT TGCCCTTCCTG TTTTGTCTCA CCCAGAAACG CTGATTAAGA
TCCTTCTCAT ACTCATAGT TGTAAAGGCA CAGCGGGAAT AAGGNNAAA ACCCGTAAA ACCGAGGAC AAAACAGAT GGTCTTTTC GACCACTTTC
3201 TAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGTTA CATCGAATCG GATCTCAACA GGGTAAAGAT CCTTGAGAT TTTGCGCCCG AAGAAAGTTT
ATTTTCTAGC ACTTCTAGC AACCCAGTG CTCACCAAT GTACCTTGC CTAGAGTTGT CCGCATTTA GGAATCTCA AAAGCGGJJC TTTCTTUCAAA
3301 TCCATGATG AGCACTTTA AGTTCTGCT ATGTGGCGG GTATTATCCC GTGATGAGC GGGCAAGAG CAATCGCTC GCGCATPACA CTATTTCTAG
AGTTACTAC TCGTGAAT TCAAGAGCA TACACCGGC CATATAGGG CACTACTGG GCCCTTCTC GTTCAGCCAG CGGCTTATCT GATTAAGACTC
3401 ATAGACTTGG TTGAGTACTC ACCAGTACA GAAGAGCATC TTAGGATGG CATGACATGA AGAGATATAT GCAGTGTGC CATAAUCATG AGTATATACA
TTACTGAACC AACTCATGAG TGGTCACTGT CTTTTCGTAG AATGCTACC GTACTGTAT TCTCTTAATA CGTCAACAGG GTATTGTATC TCACTATTTT
3501 CTGCGGCCAA CTTACTTCTG ACAACGATG GAGGACCGAA GAGCTAACG GCTTTTTCG ACAACATGGG GGATCATGTA ACTGCTTTC ATCTTTTACA
GAGCGCGT GAATGAGAC TGTGTCTAGC CTCCTGGCTT CCTCGATGG CGAAAACG TGTGTATACC CCTAGTACAT TGAGCGGAC TACCAACCTT
3601 ACCGAGCTG AATGAGCCA TACAAACGA CGAGCTGAC ACCAGTATG CAGCAGCAAT GCGCAACAG TTGCGCAAC TATTAACATG CGAATCTAT
TGGCTCTGAC TTACTTCGT ATGGTTTGT GCTCCACTG TGGTCTTAC CCGTCTCTTA CGTTGTTCG AACGCTTG ATATTTGACC SCTTGTATTA
3701 ACTCTAGCTT CCGGCAACA ATTAATAGAC TGGATGGAG CGGATAAGT TGCAGGACCA CTTCTCGCT CGGCTCTTC GGTGGCTGG TTTATTTCTG
TGAGTCCGA GGGCGTGT TAATTATCTG AACTACTCC GCCTATTTCA ACCTCTGTGT GAAGACGGA CCGGGAAGG CCGACCGACC AATTAALJAC
3801 ATAAATCTG AGCGGTGAG CGTGGGTCTC GCGTATCAT TGCAGACTG GGGCCAGATG GTAAACCTC CCGTATGTA GTTATCTACA LUALUJAC
TATTTAGACC TCGGCCACTC GCACCCAGAG CCGCATATGA ACCTGTGAC CCGGTCTTAC CATTCGGGAG GGCATACAT CAATAGATCT UCTTULUJAC
3901 TACGCACT ATGATGAC GAATAGACA GATCGCTGAG ATAGTGCCT CACTGATTA CCAATTTGTA CTGTGAGC AAGTTTACTC ATATATATTT
AGTCTGTGA TACCTACTG CTTTATCTGT CTAGGCTC TATCCAGGA GTCACTAAT COTACCAAT GACAGTCTG TTTCAATGAG TATATATJAA
4001 TAGATGAT TAAACTTCA TTTTAAAT TAAAGATCT AGGTGATCT TCCACTTCTA GGAATACPA TTAGATGAT CCAAAATCCC TTAACATGAG TTTTCTTCTC
ATCTAATCAA ATTTTGAAT AAAATTTAAA TTTTCTTCTA TCCACTTCTA GGAATACPA TTAGATGAT CCAAAATCCC TTAACATGAG TTTTCTTCTC
4101 ACTGAGCTC AGACCCGTA GAAGATCA AAGATCTTC TTGATATCT TTTTCTTC GCGTAACTG CTGCTTCAA ACAAACAAAC CACCTTATCT
TGACTCGCAG TCTGGGCT CTTTCTAGT TTCTTGAAG AACTCTAGGA AACTCTAGGA AACTCTAGGA AACTCTAGGA AACTCTAGGA AACTCTAGGA AACTCTAGGA
4201 AGCGGTGCTT TGTTCGCGG ATCAGAGCT ACCAATCTT TTTTCGAGG TACTGGCTT CAGCAGAGC CAGATACCAA ATACTGTCTT TCTACTATJAG
TCGCCCCA ACAAACGCC TAGTCTCTGA TGGTGAAG AAGGCTTCC ATTCACCGAA GTCTCTGCG GTCTATGCTT TATGACAGGA AGATCAGATC
4301 CCGTAGTATG GCGACACTT CAGACTCT GTAGACCGC CTACATACCT CCGTCTCTGA ATCTGTATC CAGTGGTGC TCGCAGTJAC GATAATCTT
GSCATCAATC CCGTGTGTA GTTCTTGA CATCTGCGG GATGATGGA GCGACACCAT TAGCAATG GTCAACAGG AGGTCACG CTATTTCAJAA
4401 GTCTTACCG GTTGGACTC AAGCATAGT TACCGATPA GCGGCGCGG TCGGGGTGAA AGCGGCTTGT GTCCACAGC CCGCTTTCG UGCTTTCJAG
CAGAAAGGCC CAACCTGAT ATCTGCTATC ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT
4501 CTACACCGAA CTGAGATACC TACAGGTGA GCATTGAA AGCGCAGCG TTCCGAGG GAGAAAGCG GAGAGTATC CCGTAAAGG CAGATTAJAA
GATGTGGCTT GACTCTATG ATCTGCTATC ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT ATGCTTAT
4601 ACAGAGAGC GCACAGGGA GCTTCAGG GGAACGCTT GATCTCTT TACTCTGTC GGGTTTCCG ACTCTGACT TGAAGTJAA TTTTCTTJAT
TGTCTCTTCG CCGTCTCTT CCAAGTCCC CTTTTCGGA CCAATAGAT ATCAGAGAG CCAAGGCG TGAAGTJAA ACTCTGACT ANAACTTJ
4701 GCTGCTCAG GGGCGGAGC CTATGAAA AGCCAGAA CCGGCTTCTT TTACGTTCC TGGCTTTCG TGGCTTTCG TGGCTTTCG TGGCTTTCG TGGCTTTCG
CGAGCAGTCC CCGCGCTCG GATACCTTT TGGCTTTCG TGGCTTTCG TGGCTTTCG TGGCTTTCG TGGCTTTCG TGGCTTTCG TGGCTTTCG TGGCTTTCG

Figure 2-3

4801 GTATATCCCT GATTCTGTGG ATACCGTAT TACGCTTT GATGAGCTG ATACCGCTCG CCGAGCCGA ACGACCGAGC GCAGCGAGTC AATTGAGCCGAG
CAATAGGGGA CTAAGACACC TATTGGCATA ATGGCGGAAA CTCACCTGAC TATGGCGAGC GCGCTCGCT TGCTGGCTCG CGTCGCTCAG TCACTCGCTC
4901 GAAGCGGAAG AGCGCCCAAT ACGCAAAACG CCTCTCCCG CGGCTTGGCC GATTCAATTAA TCCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGCA
CTTCGCCCTTC TCGCGGGTTA TCGCTTTGGC GGAGAGGGGC GCGCAACCGG CTAAGTAATT AGTTCGACCG TGCTGTCCAA AGGGCTGACC TTTCGCCCTT
5001 GTGAGCGCAA CGCAATTAAT GTGAGTTACC TCACTCAATTA GGCAACCCAG GCTTTACTT TTAAGCTTCC GGCTCGTATG TTGTGTGAA TTGTGAGCCG
CACTCGCGGT GCGTTAATTA CACTCAATGG AGTCAGTAAT CCGTGGGGTC CGAATGTGA AATACGAAG CCGAGCATAC AACACACCTT AACACTCGCC
TATTGTTAAA GTGTGCTCCTT TGTGATACT GTTACTAATG CTTAATT

>length: 5147

Figure 2-4

Figure 3
PSV.ID
length: 5171 (circular)

1 TTCAGCTCG CCGACATTTG ATTATGACT AGATCGATC GACAGCTGT GATGTGTGT CAGTTAGGT GTGGAAAGTC CCAGAGCTCC CCACAGAGCA
 AAGCTCGAGC GGGCTGTAAAC TAATAACTGA TCTCAGCTAG CTGTGACAC CTTACACACA GTCAATCCCA CACCTTTCAG GGGTCCGAGG GGTCTGTCCT
 101 GAAGTATGCA AAGCATGCAI CTCATTTAGT CAGCAACAG GTGTGGAAG TCCAGAGGT CCCAGCAGG CAGAAGTATG CAAAGCATTC ATCTCTATTA
 CTTTCATAGT TTCTGACGTA GAGTTAATCA GTCTGTGTC CACACCTTTC AGGGTCCGA GGGTCTGTC GTCTTCATAC GTTCTCTACG TAGACATTAA
 201 GTCAACCAACC ATAGTCCGC CCCTAACCTCC GCCATCCCG CCCCTAACCTC GCCCAGTTC CCCCATTCT CCCCCTCATG CCGTACTTAAI TTTTCTTTTAAI
 CACTCGTTGG TATCAGGGCG GGGATTGAG GGGTAGGCG GGGATTGAG GGGTAGGCG GGGATTGAG GGGTAGGCG GGGATTGAG GGGTAGGCG
 301 TATCAGAGG CCGAGGCGCG CTCGGCTCT GAGCTATTCC AGAAGTAGTG AGGAGGCTTT TTTGGAGGCC TAGGCTTTTG CAAAAGCTA GCTTATCTGG
 ATAGCTCTCC GGTCCGCGG GAGCCGGAGA CTCGATAAGG TCTTCATCAC TCCTCCGAAA AACCTCCGG ATCCGAAAAC GTTTTTCGAT CGAATAGGCC
 401 CCGGACCG TGCATTGGA CCGGATTCC CCGTCCCAAG AGTGACGTAA GTACCGCTTA TAGAGTCTAT AGGCCACCC CTTGGCTCTA CAGAGATATA
 GGCCCTTGGC ACCTAACCTT CCGCTTAGG GGCACGGTTC TCCTGCTATT CATGGCGAT ATCTCAGATA TCCGGTGGG GAACCGAGAT CTCTCTATAT
 501 AGCTTAGGAT TTTATCCCG GTGCCATCAT GGTTCGRCA TTAAGTCTG TCGTCCCGT GTCCCAAAAT ATGGGATTTG GCAAGAACGG AUAATTAATC
 TCGATCTCTA AATAGGGCG CACGCTAGTA CCAAGCTGGT AACTGAGGT AGCAGCGCA CAGGTTTAA TACCCCTAAC CGTCTTTGCC TCTGGATGCG
 601 TGCCCTCCG CAGGACCG GTTCAAGTAC TTCCARCAA TGACCAACAC CTCTTCAGTG GAAGTAAAC AGAATCTGT GATTATGGT AUAATAAATC
 ACGGAGGCG AGTCTTGGC CAAGTTCAIG AAGTTTCTT ACTGGTGTG GAGAGTAC CTTCCATTG TCTTAGACCA CTAATACCA TCTTTTGGCA
 701 GGTCTCCAT TCCTGAGNAG ATCGACCTT TABAGGACAG AATTAATATA GTTCTCAGTA GAGAACTCA AGAACCA CAGGAGCTC AATTTCTTTC
 CCAAGAGGA AGGACTCTT TTAGCTGGA ATTTCTCTC TTAATTATAT CAAGAGTAT CTCTGAGTT TCTTGGTGGT GCTCTCGAG TAAAAGAACG
 801 CAAAGTTTG GATGATGCT TABAGTAT TGAACACCG GAATGGCAA GTAGTGGG CATGTTTG ATAGTCGGG CCAATCTCTG TTAACAGAA
 GTTTCAAC CTACTAGGA ATTCTGAATA ACTTGTGGC CTTACCTT CATTCATCT GTACCAACC TATCAGCTC COTCAGACA AATGCTCTT
 901 GCAATGAATC AACAGGCCA CTTAGACTC TTGTGACAA GGATCATGA GGAATTTGAA AGTGACAGT TTTTCCAGA AATTGATTTG GCGAATATA
 CGGTACTTAG TTGTCCGGT GGAATCTGAG AACACTGT CTTAGTACT CTTAAACTT TCAGTGTGA AAAAGGTCT TTAACATAAC CTTTATAT
 1001 AACCTCTCCC AGAATACCA GCGTCTCTCT CTGAGGTCCA GGAGGAAA GGCATCAAGT ATAAGTTTGA AGTCTACGAG AAGAAGCT AUAAGUAGA
 TTGGAGGGG TCTTATGGT CCGCAGGAGA GACTCCAGT CTTCTTTT CCGTAGTTCA TATCAAACT TCAGATGCTC TCTTCTCTGA TTTCTCTTCT
 1101 TGCTTTCAAG TTCTCTGCT CCCTCTAAA GCTATGCAI TTTATAGAC CATGGAGCTT TTGCTGGCT TAGACCCCT TGGTCTCTT AGAALGCGAC
 ACGAAGTTT AAGAGACGAG GGGAGGATT CGATAGTAA AATATTCTG GTACCTGAA AACGACCGA ATCTGGGGA ACCGAGCAA TCTTGGGCG
 1201 TACATTAAT ACATACCTT ATGTATCATA CACATAGAT TAGTGACAC TATAGATAA CATCCACTT GCCTTCTCT CACACATCT CATTATTA
 ATGTTAATTA TGTATTGGA TACATAGTAT GTGTATCTAA ATCCACTGT ATCTTAT TATGAGTAAA CCGAAGAGA GTGTULACA CTGAGTCTCA
 1301 CAACGTCACC TCGTTCTAT CGATTGAAT CCCGGGAT CTTCTAGAT GAGCTGAC AGCTTGGCC CCAATGCCC AACTTCTTTA TTTAGCTTA
 GTTGACGTGG AGCAAGATA GCTAACTTA GGGGCCCTA GGAGATCTCA GCTGACCTC TTCGAACCG GGTACCGG TTTGAACAAAT AATCTGAAAT
 1401 TAATGTTTAC AATAAAGCA ATAGCTCAC AATTTTACA AATAAAGCAT TTTTCTACT GCATCTAGT TGTGTTTCT CCAAACTTAT CAAATTTATCT

Figure 3-1

ATTACCAATG TTTATTTCGT TATCGTAGTG TTTAAGTGT TTAATTCGTA AAAAAGTGA CGTAAGATCA ACACCAACA GGTTCAGTA GTTACATAGA
 1501 TATCATGTCT GGATGATCG GGAATTAAT CGGGGAGCA CATTGGCTG TGAAGAGGA ACTTGGTTAG GTACCTTCTG AGUUGGAAA
 ATAGTACAGA CTTAGCTAGC CTTAATTA CGCGGCTGT GGTACGGAG TTTATTGGAG ACTTCTCTT TGAACCAATC CATGGAAGAC TUCGCTCTTC
 1601 AACCACTGT GGAATGTGT TCAATAGGG TGTGAAAGT CCCCAGGCT CCAAGCAGG AGAAGTATGC AAAGCATGCA TCTCAATAG TCAGCAACAA
 TTGGTCGACA CTTTACACAC AGTCAATCCC ACACCTTCA GGGGTCCAG GGGTCTCCG TCTTCATAGC TTTCTAGT AGAGTTAATC AATCGTTTGT
 1701 GGTGTGGAJA GTCCCGAGG TCCCGAGCAG GCAAGATAT GCAAGCATG CATCTCAAT AGTCAGCAAC CATAGTCCG CCCCTAATC UCCCCATCTC
 CCACACCTTT CAGGGGTCCG AGGGGTGTC CGTCTTCAAT CGTTCTGATC GTAGAGTTAA TCAGTCTGTC GTATCAGGC GGGGATTGAG CGGGGTAGG
 1801 GCGCTAACT CCGCCAGTT CCGCCATTC TCGCCCATC TCGCCCATC GCGCTGCTA TTTTCTTAT TTATGTCAGC GCGGAGGCG CTTCCGCTTC TGAGCAATTC
 CCGGGATTGA GCGGGTCAA GCGGGTAA GCGGGGTA CCGACTGAT TAAAAAATA AATAGTCTC CGGCTCCGC GGAGCCGAG ACTCGATAAG
 1901 CAGAACTAGT GAGGAGCTT TTTTGAGG CTAAGGCTTT GCAAAAGCT GTTACCTCGA CCGGCGGCTT AATTAAGCG CGCCATTTAA ATCTGCAUG
 GTCTTCAATCA CTCCTCCGAA AAAACCTCCG GATCCGAAA CTTTTTTCGA CAATGAGCT CGCGGGGAA TTAATTCGC GCGTAAAT TAGGACUTTC
 2001 TAACAGCTTG GCACTGGCG TCGTTTACA AGTCTGTGAC TGGGAAACC CTGCGCTTAC CCACTTAAT CGCCTGAG CACATCCCC CTTCCGCCAUC
 ATTGTCCGAC CCGTACCGG ACCTAATGT TGCAGACTG ACCCTTTGG GACGCAATG CGTTGAATTA GCGGAAGCTC GTGTAGGGG GAAGCGTCTG
 2101 TGGCGTAATA GCGAAGAGG CCGCACCGAT CGCCCTTCC ACAGTTGCG TAGCCTGAAT GCGCAATGCG GCCTGATGCG GTATTCTCTC CTTACGCTATC
 ACCGCATTAT CGCTTCTCG GCGGTGGCTA GCGGGAAGG TTGTCAACGC ATCGACTTA CCGCTTACCG CGACTAGCG CATAAAGAG GAATGCTAG
 2201 TGTGCGTAT TTCACACCG ATACCTCAA GCAACATAG TAGCGGCTT GAGCGGCTT GTAGGCGCG AATTAAGCGG GCGGCTGTCG TGGTTACGCG CAGCTGATC
 ACACCCATA AGTGTGGCG TATGAGTTT CCGTGTATC ATCGCGGGA CATCGCGCG TAAATTCGC CGCCACACC ACCAATGCG GTGCGACTG
 2301 GCTACACTG CCAGCGGCT AGCGCGGCT CTTTCCCTT TCTTCCCTT AGTTCGCG AGTTCGCG GCTTCCCGG TCAAGCTCTA AATCGGCGG
 CGATGTGAAC GGTGCGGGA TCGCGGGA GGAAGCGAA AGAAGCGAG GAAAGCGG TCAAGCGG GCGGCTGTCG TGGTTACGCG CAGCTGATC
 2401 TCCCTTAGG GTTCGATTT AGTCTTTAC GGCACCTCGA CCGCAAAA CTTGATTTG GTGATGCTT ACCTAGTGG CCATCGGCTT CATAGACCTT
 AGGAAATCC CAAGGCTAAA TCACGAAATG CCGTGGAGCT GGGGTTTTT GAATAAACC CACTACCAAG TGCATCACCC GGTAGCGGGA CTATCTGCA
 2501 TTTTCGCCCT TTGACGTTG AGTCCAGCTT CTTTAATAGT GCACTCTGT TCCAAACTG AACACACTC AACCTTATCT CGGGCTATTC TTTTCAATTTA
 AAAAGCGGA AACTGCAAC TCAGTGCAA GAATATATCA CCGTGAAGA AGTTTGACC TTGTTCTGAG TTGGGTAGA GCGCGATAAG AAAACTAAAT
 2601 TAAGGATTT TSCGATTTC GCGTATTG TTAATAATG AGCTGATTA ACAAAATTT AACGCAATTT TTAACAAAT ATTAACCTTT ACAATTTTAT
 ATCCCTAAA ACGGCTAAG CCGGATAAC AATTTTATC TCGCTAAT TGTTTTAAA TTGCGCTTAA AATGTTTTA TAAATGAAA TGTAAATA
 2701 GGTGCACTCT CAGTACATC TGCTCTGATG CCGCATAGTT AAGCCACTC CGCTATCGCT ACCTGACTG GTCATGCTG CCGCGGCTGT GCGGCTGT
 CCAGTGAGA GTCATGTTAG ACGAGACTAC GCGTATCA TCCGTTGAG GCGATAGCA TTGCGTGGG CAGTACCGC GCGGCTGT GCGGCTGT
 2801 CCGCTGAG CCGCTGAG GCGTGTCTG CTCGCGGCT CCGCTGAG ACGAGCTGT ACCGTCTCG GAGCTGCTAT GTGTCAGAGG TTTTCACTAT
 GCGGACTGC GCGGACTGC CCGACAGAC GAGGCGCTA GCGAATGTC TGTTCGACAC TGGCAGAGC CCGGACTG CACAGTCTC AAAAGTGC
 2901 CATCACGAA ACGCGGAGG CAGTATCTT GAAGCAGAA GCGCTCTGT ATAGCCTAT TTTTATAGT TAAATGCTAT AATTAATAG TTTTCTTATAG
 GTAGTGCTT TCGCGCTCC GTCATAAGAA CTTCTGCTT CCGGAGCAG TATGCGGATA AAAATATCCA ATTACAGTAC TATTATTACC AAAGAATCTG
 3001 GTCAGTGGC ACTTTTCGG GAATGTGCG CGGAACCTT ATTTGTTTAT TTTTCTAAT ACATCAAT ATGTATCCG TCATGAGACA ATAACTGA
 CAGTCCACCG TGAAGAGGCC CTTTACAGC GCGTTGGGA TAAACATA AAAAGATTA TGTAGTTTA TACATAGCG AGTACTCTGT TATTGCTAT
 3101 TAAATGCTC AATAATATTG AAAAGGAG AGTATGATA TTCAACATTT CCGGTGCGC CTTATTCCT TTTTGGGCT AATTTTCTT CTTTCTTTT
 ATTACGAG TTATTATAAC TTTTCTTTC TCATACTAT AGTTGTAAA GGCACAGCG GAATAAGGGA AAAACGCG TAAACGGA GGCACAAA

Figure 3-2

3201 CTCACCCAGA AACGCTGGTG AAGTAAABAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGGGGTA AATKCTCTTGA
 GAGTGGGTCT TTGGACCAC TTTCATTTTC TACGACTTCT AGTCARCCCA CBTGCTCACC CAATGTAGCT TGACCTAGAG TTGTGCCAT TCTAGGAAT
 3301 GAGTTTTCCG CCCGAGAAC GTTTCCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG CCGGTATTTA TCCGTGTGATG ACCLUGGCA AGAGCAATC
 CTCAAAAGCG GGGCTTCTTG CAAAAGTTA CTACTCGTGA AATTTCAAG AGGATACACC GCGCCATAAT AGGGCACTAC TCGGGCCCGT TCTGTGTTGAG
 3401 GGTGCGCGCA TACACTATTC TCAGATGAC TTGTTTGTAGT ACTCACCACTG TGAGTGTGTC GTGTCTTTTC GTAGATGACC TACCGTACTG TCATTCTCTT AATACGTCAC
 CCAGCGGCT ATGTGATAAG AGTCTTACTG AACCACTCA TGAGTGTGTC GTGTCTTTTC GTAGATGACC TACCGTACTG TCATTCTCTT AATACGTCAC
 3501 CTGCCATAAC CATGAGTGAT AACACTGGG CCAACTTACT TCTGACACG ATCGGAGGAC CCAAGAGGT AACCGGTTTT TTSCACAACA TCGGGGATCA
 GACGTATTG GTACTACTA TTGTGACGCC GGTGACGCT AGCTGATGA AGCTGTGCTG GTTCCCTCGA TTGGCGAATA AACGTGTGT ACCCGTAGT
 3601 TGTAACTCGC CTGTATCGTT GGGAACCGGA GCTGAATGAA GCCATACAA AGCAGCGCG TGACACCAAG ATCGGAGCAG CAATGCAAC AATTTCTCTC
 ACATTGAGCG GAATAGCAA CCCTTGGGCT GCATTACTT CGGTATGTT TGCTGTGCTG ACTGTGTGCT TACGGTCTG GTTACCGTTG TTGCAACGCG
 3701 AACTATTAA CTGGGCACT ACTTACTCTA GCTTCCCGC AACATTTAT AGACTGGATG GAGGGGATA AGTTGCAAG ACCACTTCTG CBTCTGGGCTC
 TTTGATAAT GACCGCTTGA TGAATGAGAT CGRAGGGCG TTGTAAATTA TCTGACCTAC CTCGCGCTAT TTCAACGCTC TGTGAAGAC GCGAGCCCGG
 3801 TTCGGGCTGG CTGGTTTAT GCTGATTAAT CTGGGCTGG TGAGGCTGG TCTCGGCTA TCAATGCAAG ACTGGGCGCA GATGTAAGC CBTCTGGGCTC
 ARGCGGACC GACCAATTA CGACTATTTA GACCTCGCC ACTCGACCC AGAGGCCAT AGTAAGCTG TGACCGCGGT CTACCACTG GAGGGGCTA
 3901 CGTAGTTATC TACACGACGG GGACTCAGC AACTATGGAT GAACGAATA GACAGATCG TGAGATAGT GCTCACTGA TTAAGCAATG GTAACTCTCA
 GCATCAATAG ATGTGCTGCC CTTGATCGC TTGATACCTA CTGCTTTAT CTGTCTAGG ACTCTATCCA CGGAGTACT AATTCGTAA CATTGACAGT
 4001 GACCAAGTTT ACTATATAT ACTTTAGAT GATTAAAC TTTATTTTA ATTTAAAGG ATCTAGTGA AGATCTTTT TGATAATCTC ATCACAATAA
 CTGTTTCAA TGATATATA TGAATCTAA CTRATTTTG AATTAATTA TAAATTTCC TAGATCCACT TCTAGAAA ACTATTAGAG TACTGGTTTT
 4101 TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACC CGTAGAAC ATCAAAGAT CTCTTGAGA TCTTTTTTT CTGCGGTAA TCTGCTCTCT
 AGGATATGC ACTCAAAGC AAGTGACTC GCATCTGGG GCATCTTTTC TAGTTTCTTA GAAGAATCT AGNAATAA GACGGGCAT AGACGACGAA
 4201 GCAACAAAA AAACACCGC TACCAGCGT GGTGTTGTTG CCGATCAAG AGCTACCAAC TCTTTTCCG AAGGTACTG GCTTCAGCAG AGGCAAGATA
 CGTTTGTGTT TTTGTTGGCG ATGTGCGCA CCAACAAAC GGCCTAGTTC TCGATGGTTG AGNAATAA GCTTCATGAC CGAAGTCTG TCGCTCTAT
 4301 CCAATACTG TCCCTCTAGT GTAGCCGTAG TTAGCCACC ACTTCAAGAA CTCTGTAGA CCGCTACAT ACCTCGCTCT GCTAATCTCTG TTACCACTG
 GGTATTGAC AGGAGATCA CATCGGCATC AATCGGTG TGAAGTTCTT GAGACATCTT GCGCATCTA TCGAGCGAGA CGATTAGGAC AATGCTCACC
 4401 CTGCTGCCAG TGGGATAAG TCTGTCTTGA CCGGTTGGA CTCAGACGA TAGTTACCG ATAAAGCCCA CGGTCTGGG TGAACGGGG GTTCTGTGAC
 GACGACGGTC ACCGCTATTC AGCACAAT GGCCCACT GAGTTCTGCT ATCAATGCC TATTCCGGGT GCGCAGCCCG ACTTSCCCCG CAAGCACGTC
 4501 ACAGCCOAGC TTGGAGCGAA CGACTACAC CGACTGAGA TACCTACAG GTAGCATTTG AGAAGCCCG ACCTTTCCG AAGGAGAAA GCGGACAGG
 TGTGCGGTG AACCTGCTT GCTGGATGTG GCTTGACTCT ATGATGTG CACTCTGAC TCTTCCGGG TCGAAGGGC TTCCCTCTTT CCGCTCTCTC
 4601 TATCCGTA GCGGAGGGT CCGAACAGGA GAGGACAGA GGGAGTTCC AGGCGGAAC GCCTGGTATC TTTATAGTCC TGTGCGGT TCTCACTCT
 ATAGGCCATT GCGGTCCTA GCCTTGTCT CTGCGTCT CCGTGAAG TCCCGCTTG CCGACCATAG AATATACAG ACAGCCAAA GCGGTGAGAG
 4701 GACTTGAGCG TCGATTTTG TGATGCTCT GAGGGGCG GAGCTATG AAAACCCCA GCAACGCGC CTTTTTACG TTCTGCTCT TTTCTGCTC
 CTGAATCTG AGCTAAAAA ACTACGAGCA GTCCCCCG CTGGATACC TTTTGGGT GGTTCGCGG GAAAAATGCC AAGGACCGA AATGACCTG
 4801 TTTTGTCTAC ATGTTCTTTC CTGCTTATC CCCTGATTCT GTGATPACC GTATTACCG CTTTGTGTA -CTGATACCG CTCGCTGAG CCAATCTAAT
 AAAACGAGT TACAAGAAAG GAGCCTATG GAGCTATG CACTATTG CATATGCG GAAATCACT CAGATATGG GAGGCGCTC GATTTCTA TCTG
 4901 GAGCGCAGCG ATCTAGTGA CGAGGAGCG GAAGAGCGC CAATACGAA ACCTCTCTC CCGCGGCTT GCGGATTTA TTAATCTGAG TCGACAAATA

Figure 3-3

```
CTCGGCTCGC TCAGTCACTC GCTCCTTCGC CTCTCTGCGG GTTAIGCGTT TGGCGGAGAG GGGCGCGCAA CCGGCTAAGT AATTAGGTUG ACCGTGCTTC  
5001 GGTTCCTCCGA CTGGAAGCG GGCAGTGAGC GCACGCAAT TAATGTGAGT TACCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC TTCCGGCTTC  
CCAAAGGGCT GACCTTTCGC CGTCACTCG CGTTGGTTA ATTACACTCA ATGCAGTCAG TAATCCGTGG GGTCCGGAAT GTGAAATACG AAGGCCGAG  
5101 TATGTTGTGT GGAATTGTGA GCGGATPACA ATTCACACA GGAACAGCT ATGACCATGA TTACGAATTA A  
ATACAAACA CCTTAACACT CGCTATTGT TAAAGTGTGT CCTTGTGGA TACTGGTACT AATGCTTAAT T  
>length: 5171
```

Figure 3-4

Figure 4
PSV.IPD
length: 5712 (circular)

1 TTGAGCTCG CCGACATTTG ATTATTGACT AGATGCTATC GACGCTGTG GAATGTGTGT CAGTTAGGTT GTGGAAAGTC CCCAGGCTCC CCAACAGGCA
 AAGCTCGAGC GGGCTGTAACTA TAATAACTGA TCTCAGTAG CTGTGACAC CTTACACACA GTCAATCCCA CACTTTTCAG GGGTCCGAGG GGTGTCCTGT
 101 GAAGTATGCA AAGCATGTCAT CTCAAATTAGT CAGCAACACAG GTGTGGAAGT TCCTCCAGGCT CCCAGCAGG CAGAAGTATG CAAAGCATGTC ATCTCAATTA
 CTTTATACGT TTGCTAGTA GAGTTAATCA GTGTGTTGTC CACACCTTTC AGGGTCGGA GGGTGTGTC GTCTCTATAC GTCTCTATAC TAGAGTTAAT
 201 GTGAGCAAC ATAGTCCGCG CCGTAACTCC GCGCATCCCG CCCTAACTC CGCCCACTTC CGCCCATTTCT CCGCCCCATG GCTCACTAAT TTCTTTTATTT
 CAGTCTTGG TATCAGGGCG GGGATTGAGG CGGTAGGGC GGGATTGAG GCGGTCAAG CCGGTAAAGA GGGGGGTAC CGACTGAATTA AAAAAATAA
 301 TATGACAGG CCGAGGCCCG CTCGGCTCT GAGCTATTCC AAGTAGTG AGAGGCTTT TTGGAGGCC TAGCTTTTG CAAAAGCTA GCTTATCCUG
 ATAGTCTCC GGTCCGGCG GAGCCGAGA CTCGATAGG TCTCATCAC TCCTCGAAA AACCTCCGG ATCGAATAC GTTTTTCGAT CGATAGGUC
 401 CCGGACCG TGCATTGGA CCGGATTTC CCGTCCGAG AGTAGCTAA GTACCGCTA TAGAGCGACT AGTCCACCAT GACGATAC AAGCCACUG
 GGGCTTGC AGTAACTTT GGGCTAAG GGCACGTTT TCACTGATT CATGGCGAT ATCTGCTGA TCAGTGTGTA CTGGCTCATG TTTCGGTTC
 501 TGGCTTGC CACCCGCGAC GACGTCCCG GGGCCGTAG CCGCTTCCG GCGCGTTTC CCGACTACCC CGGCACGCG CACACGCTAG ACCTGACUG
 ACGCGAGCG GTGGCGCTG CTGCAGGGG CCGGCTATC GTGGAGCGG CCGGCAAGC GGTGTGTCG GGTGTGTCG GTGTGTCG GTGTGTCG
 601 CCACATCAG CCGGTACCG AGTGTCAAG ACTCTTCTC AGCGGCTCG GGTCTGACAT CGGCAAGTG TGGTTCGCG ACACGCTG CCGCTTCTG
 GGTGTGCTC GCGGATGCG TCGGCTTCT TGAAGAGG TCGGCGGCG CCGGCTGTA GCGTTCCAC ACCAGCGCC TGTGCGCG GCGCTTCTG
 701 GTCTGACCA CCGCGAGAG GGTGAGCGT TCGCGGCTG CCGCGGCGC ATGGCGGAGT TGAGCGGTT CCGGCTGCG CCGGCTGCG GCGCTTCTG
 CAGACTGCT GCGGCTCTC GCGCTTCTC CCGCGGCGC AGCGGCTCTA GCGGCGGCG TACCGGCTCA ACTCGGCTAG GCGCGGCGG GCGCTTCTG
 801 AGATGGAAG CCTCTGCG CCGACCGCG CCAAGGAGC CCGTGTGTT CCGGCGGCG TCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG
 TCTACCTTCC GGAGACCG GCGGTGCGG GGTCTCTCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG
 901 CCGGCTGCTG CTCCCGGAG TGGAGGCGC GAGGCGCGC GGGTGTGCG CCGTCTGGA GACTTCTGA GACTTCTGA GACTTCTGA GACTTCTGA
 GCGGAGCAG CAGGCGCTC ACCTCGCG GCTCGCGCG CCGGCGGCG CCGGCGGCG CCGGCGGCG CCGGCGGCG CCGGCGGCG CCGGCGGCG
 1001 GGTTCACCG TCACGCGCA GGTGAGTGC CCGAGGAGC GCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG
 CCGAAGTGC AGTGGGCT GCAGCTCAG GGTCTCTCG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG
 1101 TCGTGGCG GTCCCAAT ATGGGATTG CCAAGAGG GCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG
 AGCAGGCGCA CAGGCTTTA TACCCCTAAC GGTCTTGC CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG
 1201 CTCTTCTAGT GAAGTAAAC AGAATCTGT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT
 GAGAGTCTC CTCTTCTAGT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT
 1301 GTTCTCAGT GAGACTCAG AGAAGACCA CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG
 CAGAGTCTC CTCTTCTAGT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT GATTATGCT
 1401 GTAAAGTGA CAGGCTTCTG ATAGTCTG GCGTCTG GCGTCTG GCGTCTG GCGTCTG GCGTCTG GCGTCTG GCGTCTG GCGTCTG
 CATTCTAT GTACCAAC TATCAGCTC CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG CCGGAGCTG
 1501 GGAATTTGAA AGTACACGT TTTTCCGAG AATTGATTG GGAATTTG AACTCTCTCC AGAATACCA GCGGCTCTCT CTGAGTCTCA GCGGAGCTG

Figure 4-1

CCTTAAACTT TCACTGTGCA AAAGGGTCT TTAACCTAAT CCCTTTATAT TTGAGAGAGG TCTTATGGGT CCGCAGGAGA GACTCCAGGT CCTTCTTTT
 1601 GGCATCAAGT ATRAGTTTGA AGTCTACGAG AAGAAGACT AACCTTAACT GCTCCCTCTC TAAAGCTATG CATTTTATATA AGACCATGGG ACTTTTTCCTG
 CCGTAGTTCA TATTCARACT TCAGATGCTC TTCTTTCTGA TTGCATTTGA CGAGGGGAGG ATTTGATATC GTAAAAATAT TCTGGTACCC TGAAGAACGAC
 *End DHER
 1701 GCTTTAGATC CCTTGGCTT GCTTAGAGG CAGCTACAAAT TATATACATAA CCTTATGAT CATACACATA CGATTTAGGT GACACTATAG ATACATATCA
 CGAAATCTAG GGAACCGRA GCAATCTTGC GTGATGTTA ATATGTTT GGAATACATA GTATGTGTAT GCTAAATCCA CTGTGATATC TATTTAGTGT
 1801 CTTTGGCTTT CTCTCCACAG GTTCCACTC CCAGTCCAA GGTCCGGTT GAGTCCAGC CAGATAGCT AACTTAAGG GCCCTTAGGA GACCTCTTTC
 GAAACGGAAA GAGAGGTGTC CACAGTGGG GTTCCGGTT GAGTCCAGC CAGATAGCT AACTTAAGG GCCCTTAGGA GACCTCTTTC
 1901 CTTTGGCTTT CTCTCCACAG GTTCCACTC CCAGTCCAA GGTCCGGTT GAGTCCAGC CAGATAGCT AACTTAAGG GCCCTTAGGA GACCTCTTTC
 GAACTACCG CGGTACCG GTTGAACAA TAAGTCCGA TATTCACAT GTTTATTTTCG TTATCTGAT GATTTTTCAC TTTTTCCTAC
 2001 TGCATTTAG TTGTGTTTG TCCAACTCA TCAATGTATC TTATCTATG TGGATCGATC GGAATTAAT TCCGCGCAGC ACCCTGGCTT GAAATAAOCCT
 ACCTAGATC AACACCAAC AGGTTTGTAGT AGTTACATAG AATAGTACAG ACCTAGCTAG CCCTTAATTA AGCCGCTTCG TGGTACCAGA CTTTATTTGA
 2101 CTGAAGAGG AACTTGTGTTA GGTACCTTCT GAGGCGGAAA GAACCAAGT TGGATGTGT GTCTGTTAGG TCTGCAAG TCCCGAGGT CCCGAGCAGG
 GACTTTCTCC TTGAACCAAT CCATGGAAGA CTCCGCTTCT CTTGCTGAGC ACCTTACACA CAGTCAATTC CACACCTTTC AGGGCTCCGA GGGCTCTGTC
 2201 CAGAAATATG CAAGCATGC ATCTCAATTA GTCAAGAACD AGGTGTGGA AGTCCGAGG CTCCCGCAGA GGCAGAGTA TGCAGAGCT GCATCTCAAT
 GTCTTCATC GTTCTGATG TAGATTAAT CAGTCTGTG TCCACACCTT TCAGGGTTC GAGGGTCT CCGTCTTCAT ACCTTCTGTA CGTAGAGTTA
 2301 TAGTCAAGAA CCATAGTCC GCGCTTACT CCGCCATCC CCGCCCTAAC TCCGCGCAGT TCCGCGCAGT TCCGCGCAGT TCCGCGCAGT TCCGCGCAGT
 ATCAGTCTGT GGTATCAGG GGGGATTA GCGGGTATG GCGGGTATG GCGGGTATG GCGGGTATG GCGGGTATG GCGGGTATG GCGGGTATG
 2401 TTTATGAGA GCGGAGGCG GCTCGGCT CTGAGCTATT CCAGAGTAG TGAAGAGCT TTTTGGAGG CTTAGCTTT TCCGCGCAGT TCCGCGCAGT
 AATAGCTCT CCGGCTCGG GCGAGCGGA GACTCGATA GCTCTTATC ACTCTCCGA AARCTCTC GATCCGAAA ACCTTTTTCG ACAATGAGG
 2501 AGCGGCGGT TAATTAAGC GCGCATTTA AATCTGCGA GTACAGCTT GGCATGCGC GTGTTTAC AACGCTGTA CTGCGGAAA CCGGCGCAGT
 TCGCGCGGA ATTAATTCG CCGGTAAAT TTAGGACGTC CATTGTGAA CCGTGAACCG CAGCAAAATG TTGAGCAGT GACCTTTTCG GAGCGCGAAT
 2601 CCGCACTTAA TCGCCTTGA GCATATCCC CTTTCCGAG CTGGGCTAAT AGCAAGAGG CCGGACCGA TCGGCTTTC CAGCAGTTG GTAGCTTAAA
 GGGTTGANT AGCGAAGCT CTTGTAGGG GGAAGCGTC GACCGATTA TCGCTTCTC GCGGCTGCT AGCGGAAAGG GTTGTCAACG CATCGCACTT
 2701 TGGCGATGG GCGCTGATG GGTATTTCT CTTAGCCTT CTTAGCCTT CTTAGCCTT CTTAGCCTT CTTAGCCTT CTTAGCCTT CTTAGCCTT
 ACGCTTACC GCGGCTACG CCAATAAGA GGAATGCTA CACAGGCTA AATGTGTGC GTATGAGT TCGTTGAT CATCGCGCGG ACATCGCGG
 2801 CATTAAGC GCGGCTG GTGTTAGC GCAGGCTAC CCGTACACTT GCGAGGCGC TACGCGCGC TCGCTTCTC TCGCTTCTC TCGCTTCTC
 GTATTCGG CCGCCACAC CACCAATGCG CCGTACACTT CCGTACACTT CCGTACACTT CCGTACACTT CCGTACACTT CCGTACACTT CCGTACACTT
 2901 CAGCTTGGC GCGTTTCCC GTCAGCTCT AATCGGGG CTTCTTTAG GGTTCGAT TATGCTTTA CCGGCGCTC ACCCGAAA ACTTGTATTT
 GTCAAGCG CCGAAGGG CAGTTGAGA TTTAGCCCC GAGGAAATC CCAAGCTAA ATCAGAAAT CCGCTGAGG TGGGTTTTT TGAATAAA
 3001 GGTGATGTT CAGTATGG GCGATGCG TATAGAGG TTTTCCCG CTTTCCCG CTTTCCCG CTTTCCCG CTTTCCCG CTTTCCCG
 CCACTACCAA GTGATCACC CCGTACCGG ACTATCTGC AATAAGCGG AACTGCAAC CAGTACAGT TCTTATAG TGCATCTTT TCCGAAATTT
 3101 GACACACT CACCTTATC TGGGCTATT CTTTGTATT ATAGGAGT TTTCCGATTT CCGCTATTG GTTAAATAT CAGCTGATTT AAAAAAAT
 CTTGTTGTA GTTGGATAG AGCCGATAA GAAACTAAA TATTCCTAA AACGCTAA GCGGCTAA CAATTTTTTA CTCGACTAAA TTTTCTTTTA
 3201 TAACCGGAT TTTAAGAAA TATTAAGTT TACATTTTA TGTGCACTC TCAGTACAT CTGCTCTGAT GCGGCTAGT TAAGCAACT CAGCTATTTG
 ATTGCGCTTA AATTTGTTT ATAATGCAA ATGTTAAAT ACCAGGTGAG ACTCATGTTA GACGAGACTA CCGGCTATCA ATTCGCTTGA GCGCTATGCT

Figure 4-2

3301 TAGTGTACTG GGTCTATGGT GCGCCCGAC ACCGCGAC ACCGCTGAC GCGCTGTGCT GCTCCCGCA TCCGCTTACA GACAAGCTGT
ATGCACTGAC CCAGTACCGA CGCGGGGCTG TGGGCGGCTG CGCGGCTG CCGACAGAGA CAGGCGCT AGGGAATGT CTGTTCGACA

3401 GACCGTCTCC GGGAGCTGCA TGTGTACAG GTTTTCACCG TCACTACCGA AACGCGGAG GCAGTATCT TGAAGACGAA AGGCTCTCT GATACGCTTA
CTGGCAGAG CCCTCGAGCT ACACAGTCT CAAAAGTGGC AGTAGTGGT TTGCGGCTC GGTCTAAGA ACTTCTGCT TCCCGGACGA CTATGCGGAT

3501 TTTTATAGG TTAATGTCAI GATATATG GTTTCTTAGA CGTCTAGTGG CACTTTTCGG GGAATGTGC GCGGAACCC TATTTGTGTTA TTTTCTTAAA
AAAAATATCC AATTACAGTA CTATTATTAC CAAAAGATCT GCAGTCCACC GTGAABAGCC CCTTTACAGC CCGCTTGGGG ATAAACAAT AAAAAGATTT

3601 TACATTCAAA TATGTATCG CTCTAGAGC ATTAACCCCTG ATAAATGCTT CAATATATT GRAAAGGGA GAGTATGACT ATTCAACATT TCCGTGTTCG
ATGTAGTTT ATACATAGC GAGTACTCTG TTATTTGGGAC TATTTACGAA GTTATATATA CTTTTCCTT CTCTACTCA TAAGTTGTAA AGGCACAGCC

3701 CCTATTCCG TTTTTCGGG CATTTCCT TCTGTGTTT GCTCACCAG AACGCTGGT GRAAGTAAA GATGCTGAAG ATCAATTCGG TGCACGACTG
GGAATAGGG AAAAAGGCC GTAAACGGA AGGACAAA CAGTGGGTC TTTCGACCA CTTTCAATTT CTACGACTTC TAGTCAACCC ACCTGCTCAC

3801 GGTACATCG AACTGGATCT CAACAGCGT AGATCCTTG AGGTTTTCG CCGGAAAGAA CTTTTTCCAA TCACTGAGC TTTTAAAGTT CTCTATCTG
CCAATGTAGC TTGACCTAGA GTTGTCCCA TTCTAGGAA TCTCAAAAG CCGCTTCTT GCAAAAGTT ACTTCTCTG AAAATTTCA GACGATACAC

3901 GCGGGTATT ATCCGCTGAT GAGCGGGG AGAGCAACT CGGTGCGCG ATACACTATT CTCAGATGA CTTGCTGAG TACTCACCG TCACAGAAAA
CGGGCATAA TAGGCACCTA CTGGGCGCG TTCTGCTGA GCGAGCGCG TATGTATA GAGTCTTACT GAACCACTC ATGAGTGGT AGTGTCTTTT

4001 GCATCTTAG GATGCAATGA CAGTAAGAGA ATTATGAGT GCTGCTGAG TAACACTGCG GCACTTAC TCTGTGACAAC GATCGGAGCA
CGTAGAATG CTACGCTACT GTCAATCTCT TAATAGTCA CGAGGTAFT GGTACTACT ATGTGAGCG CGTTGAATG AAGACTGTG CTAGCTTCT

4101 CGAGGAGC TATCGCTTT TTTGCAAC ATGGGGATC ATGTAATCG CTTGTATCGT TGGGACCGG AGCTGAATGA AGCATACCA AACGCAAGC
GGCTTCTCG ATTTGCGHAA AAGCTGTG TACCCCTAG TACATGAGC GGAATGAGCA ACCCTGGCG TGGACTTACT TCGGTATGGT TTGCTGCTG

4201 GTGACACAC GATCGGCA GCAATGCA CAACGTGCG CAACATTA ACTGGGCA TACTTACTCT AGCTTCCCG CAACAATAA TACACTGAT
CACTGTGGT CTAGGCTGT GTTACCGT GTTGTACAG GTTGTATAT TGACCGCTG ATGAAGAGA TCGAAGGCC GTTGTAAAT ATCTGACCTA

4301 GAGGCGGAT AAGTTGCG GAGCACTCT GCGCTCGCG CTTCGGCTG GCTGGTTAT TGCTGATAA TCTGAGCGG GTGAGCGTG GTCTGCTG
CCTCCGCTA TTTCACGTC CTGCTGAGA CGGAGCGG CCGTCCGTA AGCTATAT AGACTGCG CACTGCGC CACTGCGC CAGAGCGCA

4401 ATCAATTGAG CACTGGGCG AGATGGTAA GCTCCGTA TCGTAGTAT CTACAGAGC GGGAGTCAAG CAACATGGA TGAACGAAAT AGACAGATCG
TAGTAAGTC GTGACCCCG TCTACCATC GGGGGGAT AGCTATATA GATGTCTG CCGTCACTCC GTTGTAGCT ACTTGTCTA TCTGTCTAGC

4501 CTGAGTAGG TGCTCACTG ATTAAGCAT GTTAATCTG AGACCTGCT TACTCATATA TACTTTAGAT TCAATTAAT CTTCAATTTT AATTTAAAG
GACTCTATCC ACGGATGRC TRATTGTAA CCAATTGACG TCTGTTCAA ATGAGTATAT ATCAATCTA ACTAAATTTT GAGTAAAAA TTAATTTTC

4601 GATCTAGTG AAGATCCTTT TTGTAATCT CATGACCAAA ATCCCTTAAC CACTGATTC GTTCCACTGA CCGTCAAGC CCGTCAAAA GATCAAGCA
CTAGATCCAC TTCTAGGAA ACTATTTAGA GTACTGTTT TAGGATTTG CACTCAAG CAGGTGACT CCGCTCTG GGCATCTTTT CTAGTTTCT

4701 TCTTCTGAG ATCCTTTTTT TCTGCGGTA ATCTGCTGT TGAACCAAA AAACACCG CTACAGCGG TGGTTGTTT GCGGATCAA CAGTACCAA
AGAGACTC TAGGAAAAA AGAGGCGCAT TAGAGAGA ACGTTGTTT TTTTGTGGC GATGTCCCG ACCAAACAA CCGCTAGTT CTGATGTTT

4801 CTCTTTTCC GAAGTAACT GCTTCAGCA GAGCGAGAT ACCAAATCT GTCTCTAG TGTAGCGTA GTTAGGCGC CACTTCAAG ACTCTATAC
GAGAAAAAG CTTCATTTA CCGACTCT CTGCTCTA TGCTTTTGA CAGGAGATC ACATCGGAT CATTCCGCTG GTGAAGTTCT TGAGACATCG

4901 ACCGCTACA TACCTGCTC TGTATCTCT GTTACAGTG GCTGCTGCA GTGGGATAA CTGCTGCTT ACCGCTGCT ACTCAAGAG ATATTTTAC
TGGCGATGT ATGAGGCGAG ACGATTAGA CAATGCTAC CGAGAGCGT CACCGCTAT CAGACAGAA TGGCGCAC TGAATCTGCT TATCAATGCG

5001 GATAGGCGC AGCGTCTGG CTGAACGGG GTTCTGCTCA CACAGCGCG CTTGAGCGA ACGACTTACA CCAACTGAG ATACCTACAG CTTTAACTAT

Figure 4-3

CTATTCCGGG TCGCCAGCCC GAGTTGCCCC CCAGGCAGGT GTGTGGGGT GAACTCGCT TGCTGGATGT GGCTTGACTC TATGATGTC GCACCTCGTAA
5101 GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGCGGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACACGG AGGGAGCTTC CAGGGGGA
CTCTTTCCGG GTGCGAAGGG CTTCCCTCTT TCCGCCCTGT CATAGGCCAT TCGCGCTCCC AGCCTTGTC TCTCGGTGC TCCCTCGAAG GTCCCCCTTT
5201 CGCCTGGTAT CTTTATAGTC CTGTCCGGTT TCGCCACCTC TGACTTGAGC GTCGATTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAAACCCC
GCGGACCATG GAAATATCAG GACAGCCCA AGCGTGGAG ACTGAACCTG CAGCTAAAAA CACTACGAGC AGTCCCCCG CCTCGGATAC CTTTTCGGG
5301 AGCAACGCGG CTTTTPACG GTTCTGGCC TTTTGTGGC CTTTGTCTCA CATGTTCTT CCTGGTTAT CCCCTGATTC TGTGGATAAC CGTATTACCG
TCGTTGGCC GGAAATATGC CAAGGACGG ABAACGACCG GAAACGAGT GTACAAGAA GGACCAATA GGGGACTAAG ACACCTATTG GCATAATGGC
5401 CCTTTGAGTG AGCTGATACC GCTGCCGCA GCTGCCAGC GAGTCAGTGA CGAGGAGGC GGAAGAGCG CCAATACGCA AACCGGCTCT
GGAACCTCAC TCGACTATGG CGACGGGGT CGGCTTGCTG GCTCGGTGCT CTAAGTCACT CGCTCTTGG CCTTCTCGG GGTATGCGT TTGGCGGAGA
5501 CCCCAGCGGT TGGCGATTG ATTAATCCAG CTGGCAGCAG AGGTTTCGG ACTGGAAGC GGGCAGTGA GCGAAGCAA TTAATGTAG TTACCTCACT
GGGGGGGCA ACCGGCTAAG TAATTAGTC GACCGTGTG TCCAAAGGC TGACCTTTC CCGTCACTC GCGTTGCGT RATTACACTC AATGGAGTGA
5601 CATTTGGCAC CCCAGGCTTT ACACCTTTATG CTTCCGGCTC GTATGTTGTG TGGATTGTG AGCGGATAC AATTTCACAC AGGAAACAGC TATGACCATC
GTATTCGGTG GGGTCCGAA TGTGAATATC GAAGCCCGAG CATCAACAC ACCTTAACAC TCGCCTATTG TTAAAGTGTG TCCTTTGTG ATACTGGTAC
5701 ATTACGAATT AA
TAATGCTTAA TT

>length: 5712

Figure 4-4

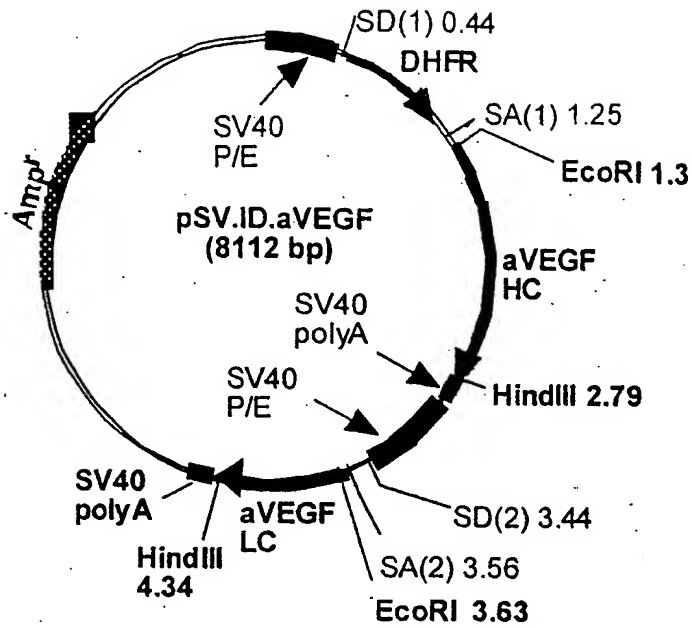


Figure 5, pSV.ID.aVEGF control plasmid

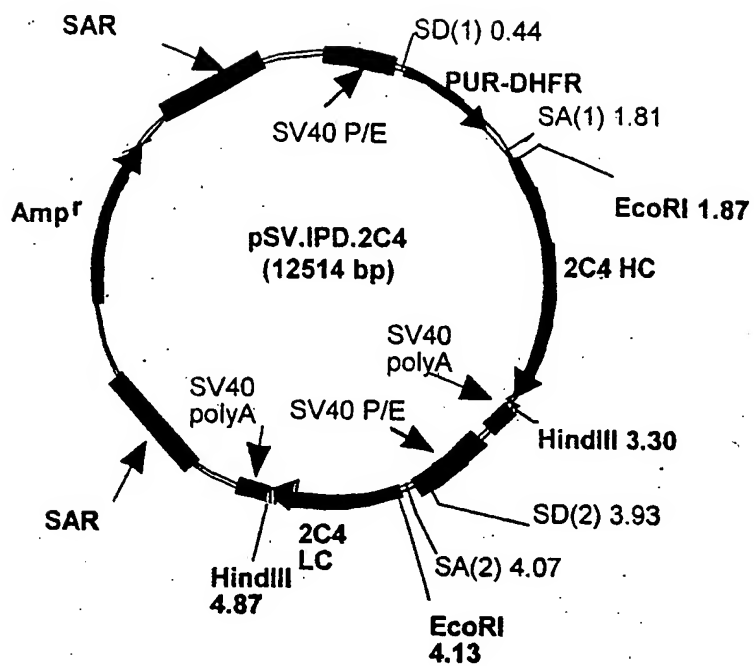


Figure 6. pSV.IPD.2C4

Figure 7
PSV.IPD.2C4
 length: 12514 (circular)

1 TTGAGCTCG CCCGACATG ATTATTGACT AGAGTGGATC GACAGCTGAG GAATGTGTGT CAGTTAGGTT GTGRRAGTC CCCAGGCTCC CCAGCAGGCA
 AAGCTCGAGC GGGCTGTATC TAATAACTGA TCTCAGCTAG CTGTGCACAC TTACACACA GTCAATCCCA CACTTTTCAG GGTCCGAGG GGTGCTCGT
 101 GAAGTATGCA AAGCATGONT CTCAATTAGT CAGCAACACAG GTGTGGAAG TCCCCAGGCT CCCACACAGG CAGAAGTATG CAAGCATGTC ATCTCAATTA
 CTTTATACGT TTCTACAGTA GAGTTAATCA GTCTGTGTC CACACCTTTC AGGGTCCGA GGGGTGCTCC GTCTTCATAC GTTTCGTAGG TAGAGTTAAT
 201 GTACACRACC ATAGTCCCGC CCCTAACTCC GGCCTACCG CCCTAACTC CCGCCAGTTTC CCGCCATTCT CCGCCCATG GCTGACTAAT TTTTITTTAT
 CAGTCGTGG TATCAGGCG GGGATTGAG GGGGTGAG GGGGTCAAG GCGGTGAAGA GCGGGGTAC CCACTGATTA AAAAAATAA
 301 TATGACAGG CCGAGGCGC CTGGCCTCT GAGCTATTCC AGAGTAGTG AGGAGCTTT TTGGAGGCC TAGGTTTTG CAAAAAGCTA GCTTATCCUG
 ATAGCTCTCC GGTCCGCGG GAGCGGAGA CTCGATAGG TCTTCATCAC TCTCCGAA AACCTCCG ATCCGAAAC GTTITTCAT CCAATAGGCG
 401 CCGGGACCG TGCATTGAA CCGGATTC CCGTCCAG AGTACGCTAA GTACCGCTTA TAGAGCGACT AGTCCACCAT GACCGAGTAC AAGCCACCG
 GGGCTTGC ACCTAACCTT GCGCTAAG GGCAGGTTT TCACTGCATT CATGGCGAT ATCTGCTGA TCAGGTGCTA TTGGGTCTAT TTGGGTGCT
 501 TGCGCTCGC CACCGCGAC GAGTCCCGC GGGCGGTAG CACCTCGCC CCGCGTTTC CCGACTACCC GCGCACGCG CACACCTAG ACCCGACCG
 ACCCGAGCG GTGGCGCTG CTGACGCGC CCGCGCATG GTGGAGCG GCGTATGG GCGGTGCGG GTGTGSCATC TTGGCTTGG
 601 CCACATCGAG CCGGTACCG AGTGCAGA ACTCTTCTC ACGCCGCTG GGTCTGACAT CGGCAAGTG TGGTTCGGG ACCAGCGCG CCGGTGCGG
 GGTGTAGCT GCGCAGTGG TCGAGCTTCT TCGAAGAG TGCGCGAG TGCGCTGCTA GCGTTCAC ACCAGCGC TCTGCGCG GCGGTGCGG
 701 GTCTGACCA CCGCGGAG GGTTCGAGCG GGGCGGTGT TCGCGAGAT CCGCGCGCG ATGCGCGGTG TGAGGGTTC CCGGTGCGG CCGCAAC
 CAGACTGTG GCGGCTCTC GCAGCTTCG CCGCGCTA AGCGCTCTA GCGGGCGG TACCGCTCA ACTCGCAAG GCGGTGCGG CCGGTGCGG
 801 AGATGGAAG CCTCTTGGC CCGCACCGC CCAAGAGCG CCGTGTGTT CTGGCCTCTC GCGCGACCG TCGCGTCTC CCGGTGCGG GTCTGCGCAG
 TCTACCTTC GGAGGACCG GCGGTGCGG GGTCTCTCG GCGACCAAG GCGCGTGGC AGCGCGAG CCGGTGCTG GTCCGTTCC CAGACCTCT
 901 CCGCTCTG CTCCCGGAG TGGAGCGCG CAGCGCGCG GGGTGGCG CCGGTGCGG GAGTCTGGA GACTTCCCG CCGGTGCGG GCGGTGCGG
 GCGGCGAC GAGGGCTC ACTCCGCG GCTCGCGG CCGCGCGG CCGGTGCGG GAGGAGCT CTGGAGCGG GCGGTGCGG GCGGTGCGG
 1001 GGTTCACG TCACGCGCA CCGTGTGTC CCGAGGAC GCGGCTCTG GCGGTGCGG GCGGTGCGG GCGGTGCGG GCGGTGCGG GCGGTGCGG
 CCGAAGTGG AGTGGGGT GCGGTCTAG GCGGTCTAG GCGGTCTAG GCGGTCTAG GCGGTCTAG GCGGTCTAG GCGGTCTAG GCGGTCTAG
 1101 TCGTGGCT GTCCCAAT ATGGGATG GCAAGACCG AGACTTACC TGCGCTCGC GCGGTGCGG GCGGTGCGG GCGGTGCGG GCGGTGCGG
 AGCAGCGCA CAGGTTTA TACCCCTAAC CCGTCTGCG CCGGTGCGG GCGGTGCGG GCGGTGCGG GCGGTGCGG GCGGTGCGG
 1201 CTCTTCAGT GAAGTAAAC AGATCTGCT GATTATGGT AGGTAACCT GGTCTCTCT TCGTGTGGA GCGGTGCGG GCGGTGCGG GCGGTGCGG
 GAGAAGTCT CTTCCATTG TCTTAGACA CTAATACCA TCTTTTGA CCAAGAGTA AGGACTCTT TTAGTGTGA ATTTCGTG CTTAATTTA
 1301 GTTCTCAGT GAGTAAAC AGACACCA CAGGAGCTC ATTTTCTTC CAAAGTTTG GATGATCTT TAAGACTTAT TGAACACCG GAATTCGCA
 CAAGAGTCT CTCTTGTG TCTTGTGT GCTCTCTAG TAAAGNAG GTTTTCAAC CTACTAGGA ATTCTGAATA ACTTGTGCG CTTAACCGT
 1401 GTAAGTAGA CATGTTTGG ATAGTCGGAG GCAGTCTCT TACACAGGA GCGATGAAT ACACAGGCA CTTAGACAT TTGTGACAA GATCTCTA
 CATTCTAT GTACCAAC TATCAGCTC GGTCAAGCA AATGTCTCT CCGTACTTAG TTGGTCCGT GGAATCTGAG AAACACTGTT CTTAGTACT
 1501 GGAATTGAA AGTACACGT TTTTCCGGA AATTGATTG GGAATATA AACCTCTCC AGAATACCA GCGGTCTCT CTGAGGTCCA CGAGGAAAA

Figure 7-1

CCTTAAACCTT TCACTGTGCA AAAAGGTCT TTAACATAAC CCTTTATAT TTGAGAGGG TCTATGSGT CCGCAGGAGA GACTCCAGGT CCTCCTTTT
 1601 GGCATCAAGT ATAAGTTTGA AGTCTAGGAG AGAAGAGACT AACGTTAACT GCTCCCTCC TAAAGCTATG CATTTTATA AGACCATGG ACTTTTGCTG
 CCGTAGTTCA TATTCAAAT TCAGATGCTC TTCTTTCTGA TTGCAATGA CGAGGGGAGG ATTTCGATAC GTAAATAAT TCTGGTACCC TGAATAACGAC
 1701 GCTTTAGATC CCGTTGGCTT CGTTAGAACG CAGCTCAAT TAATACATAA CCTTATGAT CATACACATA CGATTTAGGT GACACTATAG AATACATCC
 CGAATCTAG GGGAACCGAA GCAATCTGC GTGATGTTA ATTATGATT GGAATACATA GTATGTGTAT GCTAAATCCA CTGTGATATC TTATTTGATG
 1801 ACTTTGCCCTT TCTCTCCACA GGTGTCCACT CCGAGTCCA ACTGCACCTC GGTCTATCG ATTGAATTC ACCATGGGAT GGTCTGTAT CATCCTTTT
 TGAACGGAA ACAGAGGTGT CCACAGGTGA GGTCTCAGGT TGCCTGAGG CCAAGATAGC TAACTTAAAG TGATACCTA GTAGGAAAA
 1901 CTAGTAGCAA CTGCAACTGG AGTACATTCA GAAGTTTCAG TGCTGAGTC TGCGGTGGC CTGCTGAGC CAGGGGCTC ACTCGTTG TCTGTGCGAG
 GATCATCGTT GACGTTGACC TCATGTAAGT CTTCAAGTGC ACCACCTCAG ACGGCCAGG GACCATCTG GTCCCCGAG TGAGGCAAC AGGACACCTC
 2001 CTTCTGGCTT CACCTTACC GACTATACCA TGGACTGGGT CCGTCAAGGC CCGGTAAAG GCTTGAATG GGTTCAGAT GTTAATCTA ACAGTGGCG
 GAAGCCGAA GTGGAAAGTG CTGATATGCT ACCTGACCA GCGAGTCCG GCGCCATTC CCGACCTTAC CCAAGCTTA CAATTAGAT TGTACCCGC
 2101 CTTCTATCTAT AACAGGCT TCAAGGGCGG TTCTACTCG AGTGTGACA GATTAATAA CACATTATAC CTGCAGATGA ACAGCTGCG TGCTGAGGAC
 GAGATAGATA TTGCTCGGA AGTCCCGGC AAAGTGGAC TCACAACTGT CTAGATTTT GTCTATATG GAGTCTACT TGTGCGAGC ACGACTCTCTG
 2201 ACTGCGCTT ATTATGTGC TCGTAACCTG GGACCTCTT TCTACTTTGA CTACTGGGT CAAGGAACCC TGCTACCGT CTCTCGGC TCCACCAAGG
 TGACGGCAGA TAATAACAG AGCATTTGAC CCGTGGAGA AGATGAACT GATGACCCA GTTCTTGGG ACCAGTGGCA GAGGAGCGG AGTGTGTTCC
 2301 GCCATCGGT CTTCCCTCTG GCACCTCTT CCAAGAGCAC CTCTGGGGC ACAGCGGCC TGCGTGGCT GGTCAAGGAC TACTTCCCG AACCGGTGAC
 CCGGTAGCCA GAAGGGGAC GGTGGAGGA GGTCTCTG TGACCCCGG TGTCGCGGG ACCGACGGA CCGATTCTCTG ATGAAGGGC TTGCGCCACTG
 2401 GGTGTCTGG AACTCAGCG CCCTGACGAG CGGCTGCAC ACCTTCCCG CTGCTCTACA GTCTCTCCG CTCTACTCC TCAGCAGCGT GGTGACTGTG
 CCACAGCAC TTGAGTCCG GGTGCTGC GCGACAGTG TGAAGGGCC GACAGTGT CAGGATCTT GAGTGAAGG AGTGTGCGCA CCACTGACAC
 2501 CCTCTAGCA GCTTGGGAC CCAGACCTAC ATCTGCAAG TGAATCACA GCGGACGAC ACCAAGGTGG ACGAAGAGT TGAGCCCAA TCTTGTGACA
 GGGAGATCGT CGRACCGTG GGTCTGGATG TAGAGTTTC ACTTAGTGT CCGGTCTGT TGGTTCCACC TGTCTTCA ACTCGGTTT AGAACACTGT
 2601 AACTCACAC ATGCCACCG TGCCAGGAC CTGACTCTT GGGGGACCG TCAGTCTTCC TCTTCCCGC AAAACCCAA GACACCTCA TGATCTCCG
 TTTGAGTGTG TAGGGTGGC ACGGTCTGT GACTTGAGA CCGCTCTGC AGTCAGAGG AGNAGGGGG TTTTGGGTTT CTGTGGGAGT ACTAGAGGGC
 2701 GACCCCTGAG GTCACATGCG TGGTGGTGA CGTGAAGCAC GAAGACCTG AGGTCAAGT CACTGGGTAC GTGACGGCG TGGAGGTGCA TAATGCCAAG
 CTGGGACTC CAGTGTAGC ACCACCACT GCACTCGTG CTTCTGGAC TCCAGTTCAA GTTGACCATG CACTGCGC ACCTCCAGT ATTACGGTTC
 2801 ACAAAGCCG GGGAGGAGCA GTACAACAGC ACGTACCGG TGGTCAAGT CCTCACCTC CTGCACGAG GAGTGTTC TGACCGACTT ACCGTTCTC ATGTTCACT
 TGTTCGGG CCGTCTCTGT CATGTTGCG TGCATGCCC ACCAGTCCA GGAGTGGCAG GAGTGTTC TGACCGACTT ACCGTTCTC ATGTTCACT
 2901 AGTCTCCAA CAAAGCCCTC CCAAGCCCA TCAGAGAAC CATCTCCAA CATCTCCAA GCGAAGGGC AGCCCCGAGA ACCACAGTG TACACCTCG CCGCATCCCG
 TCCAGAGGT GTTTCGGAG GGTGGGGGT AGCTCTTTG GTAGAGTTT CCGTTCCCG TCGGGCTCT TGGTCTCCAC ATGTGGAGG GGGTAGGCG
 3001 GGAAGAGATG ACCAAGAAC AGGTACGCT GACTGCTG GTCAGAGGT TCTATCCAG CGACATGCC GTGAGTGGG AGACAATG GCAGCCGGAG
 CTTCTCTAC TGGTCTTGG TCCAGTGGGA CTGAGCGGAC CAGTTTCCGA AGATAGGCT CACTGAGG CACCTCAC TCTGTTACC CGTCGGCTC
 3101 AACACTACA AGACCAAGC TCCGTGCTG GACTCCGAG GCTCTCTCTT CTTCTACAG AGCTCACCG TGACAGAGG CAGTGGCG CAGGGGAACG
 TTGTTGATGT TGTGTTGCGG AGGCGAGC CTGAGGCTG CTGAGGCTG CAGAGAGAA GGAGATGCTG TTGAGTGGC ACCTGTTCT GTCCCTTTGC
 3201 TCTTCTCATG CTCGGTGTG CATGAGGCTC TGCACACCA CTACAGCAG AAGAGCTCT CCGTGTCTC GGTAAATGA GTGCGAGCG CCAAGAGTCC
 AAGAGATG GAGGCACTAC GACTCCGAG AGTGTGCT GATGTGGT CATGTGGT TCTCGGAGA GGCACAGAG CCAATTACT CACGCTCGG GATCTCAGC

Figure 7-2

3301 ACCTGCAGAA GCTTCGATGG CCGCATGGC CCAACTGTT TATTGCGCT TATATGGT TATATGCTT TATATGCTT TATATGCTT TATATGCTT
 TGGACGCTT CGAAGCTACC GCGGTACCG GGTGAACAA ATATACCA ATATACCA ATATACCA ATATACCA ATATACCA ATATACCA
 3401 ATTTTATTTCA CTGCATTTCTA GTGTGTTT GTCCAACTC ATCAATGAT CTTATCATGT CTGATCGGG AATTAAATCG GCGAGCACC ATGGCTGAA
 TAAAAAAGT GACGTAGAT CAACACAAA CAGTTTGG TAGTTACATA GAATAGTACA GACCTAGGCC TTAATTAAGC GCGTCGTGG TACCGGACTT
 3501 ATACCTCTG AAGAGGAAC TTGTTAGGT ACCTTCTGAG GCGGAAGAA CCAGCTGAG AATGTGTGTG AGTTAGGGTG TGGAAAGTCC CCAGGCTCCC
 TATTGGAGC TTTCTCTTG AACCAATCA TGGAGAGTC CGCTTTCTT GGTGCACAC TCAATCCAC ACCTTTCCAG GGTCCAGGG
 3601 CAGCAGCAG AAGTATGCAA AGCATGATC TCAATGATC AGCAACGAG TGTGGAAGT CCCAGGCTC CCCAGGCTC AGAATATCC AAAGCATGCA
 GTGCTCGTC TTCAATGCT TCGTACGTAG AGTTAATCAG TCGTTGTGTC ACACCTTCA GGGTCCGAG GGGTCCGAG TCTTCATACG TTTCTGACGT
 3701 TCTCAATTAG TCAGACACCA TAGTCCGCG CCTAACCTCC CCCATCCGC CCTAACCTCC GCGGATTCG GCGGATTCG GCGGATTCG
 AGAGTTAATC AGTGTGTTGT ATCAGGGCGG GGTGTAGGC GGTGTAGGC GGTGTAGGC GGTGTAGGC GGTGTAGGC GGTGTAGGC
 3801 TTTTATTTT ATGCAAGGC CGAGGCGCG TCGGCTCTG AGCTATTCCA GAAGTAGTGA GAGGCTTTT TTGAGGACT AGGCTTTTGC AAAAGCTAG
 AAAAATATA TAGTCTCG GCTCGGCG AGCGGAGAC TCGATAGGT CTTTATCAT CTTTATCAT CTTTATCAT CTTTATCAT CTTTATCAT
 3901 CTTATCGGC CGGGAACGGT GCATTGGAAC GCGGATTCG GCGGATTCG GCGGATTCG GCGGATTCG GCGGATTCG GCGGATTCG
 GAATAGGCG GCGGATTCG GCGGATTCG GCGGATTCG GCGGATTCG GCGGATTCG GCGGATTCG GCGGATTCG GCGGATTCG
 4001 TTAGAAGCG GCTACATTA ATACATAAC TTTTGGATCG ATCTACTGA CACTGATC CACTTTTCT TTTTCTCCAC AGGTGCTCC TCCCAGGTCC
 AATCTTCCG CGATGTTAT TATGTTATG AATCTTCCG AATCTTCCG AATCTTCCG AATCTTCCG AATCTTCCG AATCTTCCG AATCTTCCG
 4101 AACTGCACCT CGGTTCCGA AGTATGTTG GGTGATCG ATGAATTC ACCATGGAT GGTGATCG ATGAATTC ACCATGGAT GGTGATCG ATGAATTC
 TTGACGTGA GCGAAGGCT TCGATCGAC CCGAGTAGC TAACTAAG TCGTACCTA CCGTACCTA CCGTACCTA CCGTACCTA CCGTACCTA
 4201 AGTACATTTCA GATATCCGA TGACCCAGT ACTGGGTCT ACTGGGTCT ACTGGGTCT ACTGGGTCT ACTGGGTCT ACTGGGTCT
 TCATGTAAGT CTATAGTCT ACTGGGTCT ACTGGGTCT ACTGGGTCT ACTGGGTCT ACTGGGTCT ACTGGGTCT ACTGGGTCT
 4301 ATTGGTCTG CTTGGTATCA ACAGAACCA GGAAGGCTC GGAAGGCTC GGAAGGCTC GGAAGGCTC GGAAGGCTC GGAAGGCTC
 TRACCACAG GACCATAGT TGTCTTTGGT CTTTCTGAG CTTTCTGAG CTTTCTGAG CTTTCTGAG CTTTCTGAG CTTTCTGAG
 4401 GATCCGTTT TGGACGGAT TTCACTCTGA CCACTGCTGA CCACTGCTGA CCACTGCTGA CCACTGCTGA CCACTGCTGA CCACTGCTGA
 CTAGGCGAAG ACCCTGCTA AGTGAGACT GGTAGTCTG GGTAGTCTG GGTAGTCTG GGTAGTCTG GGTAGTCTG GGTAGTCTG
 4501 GTTGGACAG GTTACCAAG TGGAGATCRA ACGAAGTGT GGTGACCAT TGTGACAT TGTGACAT TGTGACAT TGTGACAT TGTGACAT
 CAACCTGTC CCATGTTCT ACCTCTAGT TGTGACAT TGTGACAT TGTGACAT TGTGACAT TGTGACAT TGTGACAT TGTGACAT
 4601 TCTGTTGTT GCTGCTGTA TACTTCTAT TACTTCTAT TACTTCTAT TACTTCTAT TACTTCTAT TACTTCTAT TACTTCTAT
 AGACACACA CGGACGACTT ATTGAAGATA GGTCTCTCC GGTCTCTCC GGTCTCTCC GGTCTCTCC GGTCTCTCC GGTCTCTCC
 4701 CAGAGCAGGA CAGCAAGGAC AGCATTACA GCTCAGCAG CACTTACAG CACTTACAG CACTTACAG CACTTACAG CACTTACAG
 GTCTGCTCT GTCTGCTCT GTCTGCTCT GTCTGCTCT GTCTGCTCT GTCTGCTCT GTCTGCTCT GTCTGCTCT
 4801 CCATCAGGC CTGAGTCTG CCGTACAAA GAGCTTCAAC AGGCGAGGT GTTAAAGT GTTAAAGT GTTAAAGT GTTAAAGT
 GGTAGTCCC GACTCGAG GGCAGTCTT CTGCACTCT TCCCTCTCA CACTTCAAC CACTTCAAC CACTTCAAC CACTTCAAC
 4901 TGGTTACAA TAAAGCAATA GCATCAGAA TTTTCAAAAT AAAGCATTTT TTTTCACTGA TTTTCACTGA TTTTCACTGA TTTTCACTGA
 ACCAATGTTT ATTTCTGTTT OGTAGTCTT AAAGTCTTT TTTTCACTGA TTTTCACTGA TTTTCACTGA TTTTCACTGA
 5001 CATGCTGGA TCGGGATTA ATTGCGGCA GCACATGAG CTGAATAAG TTTTAAACCT CTGAATAAG TTTTAAACCT CTGAATAAG
 GTACAGACCT AGCCCTAAT TTAGCCCGT CTTGTTGGA GACTTTCTCC TTTGAACCAAT CCATGGCTGA TCATGTTCC

Figure 7-3

5101 TCGCCACGCA CAAGATCAAT ATTRACATC AGTCATCTCT CTTTAGCAAT AAAAGGTGA. AAATATACAT TTTAAAATG ACACCATAGA CGATGTATCA
 AGCGTGCGT GTTCTAGTTA TAATGTTAG TCAGTAGAGA GAATCGTTA TTTTCCACT TTTTAATGA AATTTTAC TGTGSTATCT GCTACATACAT
 5201 AAATAAATCTA CTTGGAATGA AATCTAGGCA AGAAGTGCAG AGACTGTTAC CCAGAAAACCT TACAAAATTGT AAATGAGAGG TTAGTGAGAG TTTAAATGAA
 TTTATTAGAT GAACCTTTAT TTAGATCGT TCTTCACCT TCTGACAAAG GGTCTTTGA ATGTTTAACTA TTTACTCTCC AATCACTTCT AAATTTACTT
 5301 TGAAGATCTA AATAACTTA TAAATGTGA GAGAAATTA TGAATGTGA AGTTAATGCA GAAACGAGAG GACATACTAT ATTCATGAAC TAAAGACTT
 ACTTCTAGAT TTAATTTGAAT ATTAACACT CTCCTTAAT ACTTACGAT TCAATTTACGT CTTTGCCTCT CTGTATGATA TAAGTACTTG ATTTCTGAA
 5401 AATATTGGA AGGTATCTT TCTTTTACA TAAATTTGA GTCAATATCT TCACCCCAA AAAGCTGTTT GTTAATCTGT CAACCTCTT TCAAAATGTA
 TTATAACACT TCCATATGAA AGAAAGTGT ATTHAACAT CAGTTATACA AGTGGGTTT TTTGACAAA CAATGACAA GTTGGAGTAA AGTTTATCAT
 5501 TATAGAAAGC CCAAGACAA TAACAAATAT ATCTTCTAG AACAAATCG GAAAGATCT TCCACTAAAT ATCAAGATTT AGAGCAAGC ATGAGATGTG
 ATATCTTTTC GGTTCCTGT ATTGTTTTTA TAAGAACATC TTGTTTTAC CTTTCTTACA AGGTGATTTA TGTCTTAAA TCTCGTTTCG TACTCTACAC
 5601 TGGGGATAGA CAGTGAGCT GATAAATAG AGTAGAGCTC AGAAACAGAC CCAATGATAT ATGTAAGTGA CCAATGAAA AATATGGCA TTTTACAATG
 ACCCTATCT GTCACTCGA CTATTTATC TCTCTCGAG TCTTGTCTG GCTAACTADA TACATTCACCT GGTACTTTT TTTATACCGT AATATGTTAC
 5701 GGAATATGAT GATCTTTTC TTTTATAGAA AACHGGGAA ATATATTTAT ATGTAAAAA TAAAGGGAA CCAATATGTC ATACCATACA CACAAAAAAA
 CCTTTTACTA CTAGAAAAG AAAAAATCT TTTGTCCCTT TATATAATA TACATTTT ATTTCCCTT GGTATATACAG TATGTTATGT GTCTTTTTT
 5801 TTCCAGTGAA TTATAAGTCT AAATGGAGAA GGCAAACTT TAAATCTTTT AGAAATAT ATAGAGCAT GCCATCATGA CTTAGTGTGA GAGAAAAAT
 AAGTCACTT AATATTCAGA TTTACCTCTT CCGTTTGA AATTAGAAAA TCTTTTATTA TATCTCTGA CGGTAGTACT GAATGACAT CTCTTTTAA
 5901 TCTTATGACT CAAAGTCTA ACCACAAAGA AAGATCTCTT AATTAGATG CATGAATAT AAGACTTAT TTTAAATTA AAAAACCATT AAGAAAAATC
 AGATCTGA GTTTCAGAT TGGTGTCT TTTCTACAA TTAATCTAAC GTCTTATAA TCTGATAA AAATTTTAA TTTTGGTAA TTTCTTTTCAG
 6001 AGGCACTAGA ATGACAGAA ATATTTGCA CACCCAGTA AAGAGATG TAATATGAG AATATAAAA GAAGCTTAC AATCAGTAA AAAAAAATC
 TCGGTATCT TACTGTCTT TATAACGCT GTGGGCTCT TCTCTTAAC ATATACGTC TAATATTTT CTTCAAGATG TTAGTCAAT TTTTATTTG
 6101 TAGACAAAA TTTGACAGA TGAAGAGAA ACTTAATA ATCATCAC ATGAAACT CAATCTCAGA AATCAGAGAA CTATCTGC ATATACACTA
 ATCTGTTTT AAATCTCTCT ACTTCTCTT TGAGATTTAT TAGTAATGT TACTCTTGA GTTAGATCT TTAGTCTCTT GATAGTAAG TATATGTAT
 6201 AATTAGAGAA ATATTAAG GCTAAGTAC ATCTGTGCA ATATTGATG TATTAACCT TGATATGATG TGATGAGAAC AGTACTTTAC CCAATGGCT
 TTAATCTCT TATAATTTT CATTATCTG TAGACACCT TATACTTACC ATATATTGA ACTATCTAC ACTACTCTTG TCATGAAATG GGTACCCCG
 6301 TCTCCCCAA ACCCTTACC CAGTATAAAT CATGACAAAT ATACTTTAA AACCATTACC CTATATCTAA CCAGTACTCC TCAAACTGT CAAGTCTATC
 AGGAGGGGT TGGGAATGG GTCATATTTA GTACTCTTTA TATGAATTT TTGTTATGG GATATAGAT GGTCTGAGG AGTTTGAACA GTTCCAGTAG
 6401 AAAAAAAGA AAGTCTGAG GAATCTGCA AACTAAGAGG AACCAGAGG AAGATGAGAA TTTATATGTA TGTGGCATTC TGAATGAGT CCCAGAACAG
 TTTTATCT TTTCACTC CTTGACATG TTGATCTCC TTGGTCTCT CTGTACTCTT AATATACAT ACACCGTAAG ACTTACTCTA GGTCTTGTCT
 6501 AAAAAAGACA GTAGCTAAA AACTAATGAA ATATAATAA AGTTTGAAT TTAGTTTTT TTAABAAGA GTAGCATAA CACGGGAAAG TCATTTTCTAT
 TTTTCTGT CATCGATTT TTGATTAAT TATATTAAT TCAAACTG AATCAAAA AAATTTTCT CATGTAAT GTGCCGTTT AGTAAAGTA
 6601 ATTTTCTTG ACATTAAGT ACAAGTCTAT AATTAATAA TTTTAAATG TACTCTGAA CATTGCCAGA AACAGAGTA CAGCAGCTAT CTGTGCTCTC
 TAAAAAGAC TTGTAATCA TGTTCAGTA TTAATTTTA AAAAAATTTT ATCAAGCTT GTAACTGCT TGTCTTCTAT GTCGTGATA GACAGCAGAG
 6701 GCCTAACTAT CCATAGCTGA TTGGTCTAAA ATGAGATACA TCAACGCTCC TCAATGTTTT TTGTTTTCTT TTTAAATGAA AAATTTTAT TTTTAAAGAG
 CGGATTGATA GSTATCGACT AACCAGATTT TACTCTATGT AGTTGCGAGG AGTACAAAA AACAAAGAA AATTTACTT TTTGATATAA AAATTTCTC
 6801 AGTTTCAGT TCATAGCAAA ATTAGAGGA AGGTACATTC AAGCTCAGGA AGTTTCTCT TATCTCTAGT TTAGTGAGAG ATTGCATCAT GAATGGGTCT

Figure 7-4

TCAAGTCCA AGTATCGTTT TAACTCTCT TCCATGTAG TCGACTCT TCAAAAGGAG ATAAGATCA AATGACTCTC TAACGTAGTA CTTACCACACA
 6901 TAAATTTTGT CAAATGCTTT TCTGTGTCT ATCAATATGA CAGTGTGAT TCTTCTTTTA ACCTGTTGAT GGCACAAAT ACCTTAAATG ATTTTCAAC
 ATTAAACA GTTTACGAAA AAGACACAGA TAGTTATACT GGTACACTAA AAGAGAAAT TGGACAACTA CCTGTGTTAA TGCATTAAC TAAAGTTTG
 7001 GTTGAACAC CCTACATAT CTGGATAAA TTCTACTTGG TTGGGTGTA TATTTTGA TACATCTTG GATCTTTTT GCTAATATTT TGTGAAAT
 CAACTTGGT GGAATGTATA GACCTTATTT AAGATGAACC AACCCACAT ATAAABACT ATGTAGAAC CTAAGAAAA GATTATATA ACAACTTTTA
 7101 GTTGTATCT TTGTATGA GAGATATGG TCTGTGTTT TCTTTCTG TAAATCAT TCTAGTTCC GGTATPAAG TAAATGTGG CTAGTTGAAT
 CAAACATAGA AACAAGTACT CTCTATAACC AGACAACAA AAGAAAGAC ATTACAGTAA AAGATCAAGG CCATAATTC ATTACGAAG GATCAACTTA
 7201 GATTTAGAA GTATTCCTC TGTCTGTG TTCTGAGTA CCGCGCGC CCGTGTGTT ACAAGTCTGT GACTGGGAAA ACCCTGGCGT TACGGAACCT
 CTAATCCTT CATAGGGAG AGRAGACAG AGRATCCAT GCGCGCGG GGCACAAA TGTTCAGCA CTGACCTTT TGGGACCGCA ATGGGTTGTA
 7301 AATCGCCTT CAGCACATCC CCTTTGCG AGCTGGGTA ATAGGAGCA GCGCGCGC GATGCGCTT CCAACAGTT GCGCAGCTG AATGGCGGAT
 TTAGCGGAC GTGTGTAGG GGAAGGCG TCGACCGCAT TATCGTTCT CCGGGCGTGG CTAGCGGAA GGTGTGTCAA CCGTCTGGAC TTACCGCTTA
 7401 GCGCCTGAT GCGTATTTT CTCTTAGC ATCTGTGG TATTCACAC GGCATAGTC HAAGCAACCA TGTAGCGCG CCTGTAGCG CCGATTAGC
 CCGCGACTA CGCCTAAA GAGGATGCG TAGACAGCG ATAAATGTG GGTATGTCG TTTGTTGT ATCATGCGG GGCATCGCC GCGTATTCG
 7501 GCGCGGGTG TGTGTTTAC GCGCAGGTG AGCGTACAC TTGCGAGCG CCTAGCGCC GCTCTTTG GCTCTTTCC CACACGTTG GCGACGTTG
 GCGCGCCAC ACCACCAATG GCGTGGCAG TGGCATGTG AACGTGCG GATCGCGG GAGGAAAGG AAGGAAAGG CCGTGGCAGG
 7601 CCGCTTTCC CCGTCAAGCT CTAAATGCG GGTCTCTTT AGGTTTCCA TTAGTGTCT TACGCACTT CGACCCCAA AAATTTGAT TGGTGTATG
 GCGCGAAGG GCGAGTTGA GATTTAGCC CCGAGGAAA TCCCRAGGT AAATCAAGAA ATGCCGTGGA GCTGGGTTT TTTGAACATA ACCCACTAC
 7701 TTCAGTAGT GCGCATCGC CTGTATAGC GGTTTTCC CTTTGAAGT TGGAGTCCAC GTCTTTTAT GTTCTCAAA TGTTCGAAAC TGGAAACACA
 AAGTGATCA CCGGTAGCG GACTATCTG CCAAAAGCG GGAATGCA ACCTCAGGT CHAGAAATTA TCRCTGAGA ACAAGTTTG ACCTTTGTT
 7801 CTCACCCCTA TCTCGGCTA TCTTTTGT TATAAGGA TTTTCCGT TCTGCTAT TGGTAAAAA ATGAGTGT TACTCGACTA AATGTTTTT AATTTGCGT
 GAGTTGGAT AGACCCGAT AAGAAACTA AATATTCCT AAACGGCTA AAGCGGATA ACCAATTTT TACTCGACTA AATGTTTTT AATTTGCGT
 7901 ATTTTACAA AATATTACG TTTACATTT TATGTTGAC TCTAGTACA ATCTGCTG ATGCCGATA GTTAAGCCAG CCGCGACAC CCGGTTGTG
 TAAATTTGT TTATATTTG AANTGTTAA ATCCACGTG AAGTCTGT TAGACGAGAC TACGGGTAT CAATTCGGT CCGGTTGTG
 8001 CCGTGACCG CCGTACCGG CTGTCTGT CCGCATCC GGTACAGC AAGTGTGAC GGTCTCGG AGCTGATGT GTACAGGTT TACACCGTCA
 GCGACTGCG GCGACTGCC GAACAGCA GCGCGTAGS GAAATGTCTG TCGACACTG CCGAGGCGG TCGAGGTACA CAGTCTCAA AAGTGGCAGT
 8101 TCACCGAAC GCGCGRAGA CGAAAGGCC TCGTATAGC CCTATTTTA TAGTTAATG TCATGATAAT AATGTTTCT TAGACGTGAG GTGGCATT
 AGTGCTTTG CCGCTCTCT GCTTTCCCG AGCACTATG GGTAAATAT ATCCAATTAG ATACTATTA TTACCAAGA ATCTGCAGT CACCGTGA
 8201 TCGGGAAAT GTGCGCGAA CCCCTATTTG TTTATTTTC TAAATACAT CAAATATGTA TCCGCTCATG AGACATAAC CCGTGAATAT GCTTCAATAA
 AGCCCTTTA CACGCGCTT GGGGTAAC AATATAAAG ATTTATGTA GTTTATACAT AGGCGAGTAC TCTGTTATG GACTATTTA CGAAGTTAT
 8301 TATTGAAAA GGAAGATAT GAGTATCAA CATTTCCG TCGCGTAT TCCCTTTT GCGCATTTT GCTTCTCTAC CAGAAACCG
 ATAACTTTT CCTTCTATA CTATPAAGT GTAAAGGAC ACGGGAATA AGGAAAAA GCGCGTAAA CCGAGGACA AAAACGAGT GGTCTTTGCG
 8401 TGGTGAAGT AAAAGATGCT GAAGATGCT TGGTGGCAG AGTGGTTAC ATCTCAGAC CCGTAAATG CTTGAGATG TTGCGCTGAG
 ACCACTTTA TTTTCTAGCA CTTCTAGTCA ACCACGTGC TCACCAATG TAGCTTGGC GCAATCTAG GAACTCTCA AAGCGGCT
 8501 AGRACGTTT CCAATGATGA GCATTTTAA AGTTCTGCTA TGTGCGCGG TATTAATCCG TATTACGCG GCGGAGAGC AACTCGGTG CCGCATACAC
 TCTTGCAAA GGTACTACT CCGTAAAT TCAAGACAT ACACGCGG ATAAAGGCG ATACTGCG CCGTCTCTG TTGAGCCAG GCGGTATG

Figure 7-5

8601 TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAAATTATG CAGTCTCTGCC ATAACCATGA
 ATAAGAGTCT TACTGACCA ACTCATAGT GGTCAATGTC TTTTCGTAGA ATGCTTACCG TACTGTCAAT CTCCTAATAC GTCACGCGG TATTGTACT
 8701 GTGATAACAC TGGGGCCAC TTACTTCTGA CACGATCGG AGGACCGAG GAGCTAACCG CTTTCTTGA CACATGCGG GATCATGTAA CTCGCCCTTGA
 CACTATTGTG ACGCGGTG AATGAAGACT GTTGCTAGCC TCTGGCTTC CTGAAATGGC GAAATAAGGT GTTGATACCC CTAGTACATT GAGCGGAAT
 8801 TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAGCAG CAGGCTGACA CCAGATGCC TGTAGCAATG GCACAAAGT TGGGAAACT ATTAAGTGGC
 AGCAACCCCTT GGCCTGAGT TACTTCGGA TGGTTGCTG CTGCACTGT GGTCTAGG ACATCGTTAC CGTTGTTGA ACGGTTTGA TAATTGACCG
 8901 GAACACTTGA CTCTAGCTTC CCGGCAACAA TTAATAGACT GATGAGGAG GATTAAGTT GCAGACAC TCTGCGCTC GGCCTTCCG GGTGGTGGT
 CTTGATGAT GAGATCGAG GCGCGTTGTT AATTATCTGA CCTACCTCG CTAATTTCAA CGTCTGCTG AAGAGCGAG CCGGAGGCG CGACCGACCA
 9001 TTAATGCTGA TAAATCTGA GCGGTGAGC GTGGGCTCG CGGTATCATT GCAGACTGG GGCAGATGG TAAGCCCTCC GGTATCGTAG TTATCTACAC
 AATAAGACT ATTAGACCT CGGCGACTCG CACCCAGTCC GGCATAGTAA CGTGTGACC CCGTCTACC ATTGCGGAGG GCATAGCATC ATAGATGTG
 9101 GACGGGAGT CAGGCAACTA TGGATGAACG AATAGACAG ATCGCTGAGA TAGTGCTC TAGTGCTC ACTGATTAG CATTTGGTAAC TGTGAGACCA AGTTTACTCA
 CTGCCCTCA GTCCGTGTGAT ACTACTTGC TTTATCTGTC TAGGACTCT ATCCAGGAG TGACTAATTC GTAACTATG ACAGTCTGT TCAAATGAGT
 9201 TATATACTTT AGATTGATTT ABACTTCAT TTTTAAATTA AAGGATCTA GGTGAAGTC CTTTGTGATA ATCTCATGAC CAAATCCCT TAACGTGAGT
 ATATATGAA TCTACTTAA TTTTGAATTA AATTAATTA TTTCTAGAT CCACCTTAGT GAAATACTAT TAGAGTACTG GTTTTAGGGA ATTGCACTCA
 9301 TTTCTGTTCA CTGAGCGTCA GACCCGTAG AAGATATCAA AGGATCTCT TGAGATCTCT TTTTCTGCG CGTAATCTG TGTGTCANA CAAAAAACC
 AAGCAAGGT GACTCGCAT TTTCTAGT TCTTAGTAA ACTTAGTAA TTTCTAGT TCTTAGTAA AAAAGAGCG GCATAGACG ACGAACGTT GTTTTCTG
 9401 ACCGCTACCA GCGTGGTGT GTTTGCGGA TCAAGACTA CCACTCTTT TTCCGAGGT AACTGGCTC AGCAGAGCG AGATACCANA TACTGTCTCT
 TGGCGATGT CGCCACCAAA CAAACGCT AGTCTGAT GTTTGANA AAGCTTCCA TTGACCGAG TCGTCTGCG TCTATGGTT ATGACAAGAA
 9501 CTAGTGTAGC CGTAGTTAG CCACACTTC AAGACTCTG TAGCACGCG TACTACTCT GCTCTGTAA TCTGTCTACC AGTCTCTCT GCCAGTGGC
 GATCATCG GCATCAATCC GTTGTGAG TTCTTGAGAC ATCGTGCGG ATGTATGAG CGAGACGAT AGGACATGG TCACCGAGA CGGTCAACCG
 9601 TAAAGCTGT TCTTACCGG TTGGACTCA GAGATAGT ACGGATAAG GCGCAGCGT CCGCTGAAC GGGGGTTCG TGCACACAGC CCAGCTTGA
 TATTAGCAC AGAATGGCC ACCTGAGT CTGCTATCAA TGCCCTATTC CCGCTGCCA CCGCGACTTG CCGCCCAAGC ACGTGTGTG GGTCAACCT
 9701 GCGAAGACC TACACCGAC TGAGATACCT ACAGCTGAG CTATGANA GCGCCACCT TCCCGAAGG ABAAGCGG ACAGGTATCC GGTAGCGCG
 CGCTTGTG ATGTGCTTG ACTCTATGA TGTGCACTC GATCTCTTT CGCGGTGCGA AGGCTTCCC TCTTCCGCG TGTCATAGG CCATTGCGCG
 9801 AGGTCGGAA CAGGAGCG CAGGAGGAG CTTCAGGGG GAAACGCTG GTATCTTTAT AGTCTGTGCG GTTTCGCCA CCTCTGACTT GAGGTCTGAT
 TCCAGCCCTT GTCTCTGCG GTGCTCCCTC GAAGTCCCC CTTTGGGAC CATAGANA TCAGGACAGC CCAAGCGGT GAGACTGAA CTCGAGCTA
 9901 TTTTGTGATG CTCGTACGG GGGCGAGCC TATGANAATA GCGCGCTTTT TACGGTCTT TACGGTCTT TACGGTCTT TACGGTCTT TACGGTCTT
 NAACACTAC GAGCAGTCC CCGGCTCGG ATACCTTTT GCGGTGTTG CCGCGGAAA ATGCCAAGGA CCGGAAAAG ACGGAAAAG GAGTGTACAA
 10001 CTTTCTGCG TTATCCCTG ATTCTGTGA TAACTGAT TACCGTAT TACCGTATG AGTACTGA TACCGTCTG CCGAGCGGAA CCGAGCTGCA CAGGAGTCA
 GAAAGAGCG NATAGGGAC TAAGACACT ATTGGATRA TGGCGAATC TCACTGACT ATGGCGAGC GGTCTGCTG GGTCTGCTG GGTCTGCTG
 10101 GTGAGGAGG AAGCGGAAGA GCGCGCGG CAGGCTGCCA CCGCAAGAT CAATTTAAC AATCACTCAT CTCTCTTAG CAATAAAG GTGAAAAAT
 CACTGCTCC TTGCGCTTCT CCGGCGCCG TTCCAGCGGT GGTGTCTA GTTATATG TTAGTCACTA GAGAGATC GTTATTTTC CACTTTTAA
 10201 ACATTTTAA AATGACCA TAGAGATGT ATGAATAA TCTACTTGA AATATCTA GGCAGAGAG TGCAGACTG TTACCCAGAA AACTTACAAA
 TGTAAATTT TACTGTGT ATCTGTCTA TACTTTAT AGATGAACCT TTTTATAGT CCGTTCTTC ACGTCTGAC AATGGTCTT TTGAATGTT
 10301 TTGTAATGA GAGGTAGT AAGATTTAA TGAAGAAGA TCTAATRAA CTTAATAT GTGAGAGAA TTAATGATG TCTAAGTTAA TGCAGAAACG
 AACATTTACT CTCATACAC TTCTAATTT ACTTACTCT AGATTTATTT GAATATTTA CACTCTCTT AATTAATAC AGATTTCAAT ACCTCTCTG

Figure 7-6

10401 GAGAGACATA CTATATTTCAT GAACTAABAG ACTTAATATT GTGAGGTAT ACTTCTCTTT CACATAAATT TGTAGTCAAT ATGTTACACC CAAAAAGCT
CTCTCTGTAT GATATTAAGTA CTGATTTTC TGAATTATA CACTTCCATA TGAAGAAAA GTGATTTTAA ACATCAGTTA TACAAGTGG GTTTTTTCGA
10501 GTTTGTTAAC TTGTCACCT CATTTCAAAA TGTATATAGA AGCCCAAG ACAATAACAA AATATTCTTT GTAGAACAAA ATGGGAAGA ATGTTCCACT
CAAACAATG ACAGTTGGA GTAAAGTTT ACATATATCT TTGGGTTTC TGTATTGTT TTTATAAGAA CATCTGTTT TACCTTTCT TACAAGGTGA
10601 AATATCAAG ATTTAGGCA AAGCATGAGA TGTGTGGGA TAGACAGTA GGCTGATAAA ATAGAGTAGA GCTCAGAAC AGACCCATG ATATATGTAA
TTTATAGTTC TAAATCTGT TTCTACTCT ACACACCT ATCTGTCACT CCGACTATTT TATCTCATCT CGAGTCTTG TCTGGTAA TATATACATT
10701 GTGACCTATG AAAAAAAT GGCATTTTAC ANTGGAAAA TGTATCTTT ACTACTAGAA RAAGAAAAA TCTTTTCTC CTTTATATA AATATACATT TTTTATTTT
CACTGGATAC TTTTTTATA CCGTAAATG TTACCCCTTT ACTACTAGAA RAAGAAAAA TCTTTTCTC CTTTATATA AATATACATT TTTTATTTT
10801 GGAACCCATA TGTCAATACA TACACACAAA AAAATTCAG TGAATTATA GTCTAATG AGAGGCAAA ACTTTAAATC TTTTAGAAAA TAATATAGAA
CCTTGGGTAT ACAGTATGTT ATCTGTCTTT TTTTAGGTC ACTTAATATT CAGATTTACC TCTTCCGTTT TGAATTTAG AATCTCTTT ATTATATCTT
10901 GCATGCCATC ATGACTTCAG TGTAGAGAAA AATTTCTTAT GACTCAAGT CCTAACACA AAGAAAGAT TGTATATTAG ATTGCATGAA TATTAAGACT
CGTACGGTAG TACTGAAGTC ACATCTCTTT TTAAGATA CTGAGTTCA GGATGGGT TCTTTTCTA ACATTAATC TAACGTACTT AATATCTGA
11001 TATTTTAAA ATTAAAAAC CATTAGAAA AGTCAGGCCA TAGATGACA GAAATATTT GCAACCCCC AGTAACAGAA ATTGTAATAT GCAGATTATA
ATAAAAAAT TAAATTTTG GTAAATCTTT TCAGTCCGT ATCTACTGT CTTTATATA CGTTGGGG TCAATCTCT TACATTAATA CGTCTAATAT
11101 AABAGAGTC TTACAATCA GTAAAAATA AACTAGACA AAAATTTGAA CAGATGAAG ARAACTCTA AATATCAAT ACATAGAGA AACTCAATCT
TTTTCTCAG AATGTTTAT CATTTTTAT TTTGNTCT TTTTAACTT GTCTACTTC TCTTGAGT TTAATAGTAA TGTGTACTCT TTGAGTAGA
11201 CAGAAATCAG AAGCTATCA TTGCATATAC ACTAANTTAG AGAANTTA AAGGCTAAG TAACATCTGT GGCAATATG ATGATATATA ACCTTGATAT
GTCTTTAGTC TCTTGATAGT AAGGTATATG TGAATTAATC TCTTATAAT TTTCCGATTC ATTGTAGACA CGTTATAAC TACCTATAT TGGAACTATA
11301 GATGTATGA GAACGTACT TTACCCATG GGTCTCTCC CCAACCCCTT ACCCCAGTAT AATATCACTT TAAARCCAT TACCCATAT
CTACACTACT CTGTCTATGA ATGGGGTAC CCGAAGGAG GTTTGGGAA TGGGTCTATA TTTAGTACTG TTTATATGAA ATTTTGGTA ATGGATATA
11401 CTAAACAGTA CTCTCAAAA CTGTCAAGT CATCAAAAT AAGAAAGTC TGAGGACTG TCAAACTAA GAGGAACCA AGGAGCATG AGAATATAT
GATTTGTCAT GAGGATTTT GACAGTTCCA GTAGTTTAA TTCTTTTTCAG ACTCTCTGAC AGTTTCTGAT CTCTTGGGT TCTCTCTGAC TCTTAATATA
11501 GAAATGTGC ATTCTGAATG AGATCCAGA ACAGAAAG ACAGTAGCT ARAACTATA TGAATATATA ATRAGTTTG AACTTTAGT TTTTATATA
CAATACACG TAAGACTTAC TTAGGGTCT TGTCTTTTC TTCTATCGA TTTTGGAT ACTTATAT TATTTCAAC TTGAATCAA AAAAAATTTT
11601 AAGGTAGCA TTAACACGC AAGTCAATTT TCATATTTT CTGACATTT AATGTAATTA AATTTTTTA AATGTAGTCT GGAATCTCC
TTCTCATCT AATGTGCGG TTTCAGTAA AGTATAAAA GACITTTGTA TTCTATGTTA GATATTAAT TTTAAAAAT TTACATCAGA CCTTGTAAAG
11701 CAGAACAGA AGTACAGCAG CTATCTGTGC TGTGCGCTAA CTATCCATG CTGATGGTC TAAATGAGA TACATCAAG CTCCTCATG TTTTGTGTTT
GTCTTTGCT TCATGTGCTC GATAGACAG ACAGCGAT GATAGTATC GACTAACCG AATTTACTCT ATGTAGTTGC GAGGAGTAC AAAAAACAAA
11801 TCTTTTAAA TGAACAACT TATTTTAAA GAGGATTC AGTTCTATG TCCCTCAAG TCCCTCAAG TCCCTCAAG TCCCTCAAG TCCCTCAAG
AGAAAAATTT ACTTTTGA AAAAAAAT CTCTCAAG TCCCTCAAG TCCCTCAAG TCCCTCAAG TCCCTCAAG TCCCTCAAG TCCCTCAAG TCCCTCAAG
11901 TAGTTTACTG AGGATTTGCA TCAATGAATG GTGTTAATTT TGTCAATG CTTTTCTGT GCTATCAAT ATGACCATG GATTTCTTC TTTAACCTGT
ATCAATGAC TCTCTACGT AGTCTTACC CACAATTTA RACGTTTAC GAAAAACA CAGATAGTTA TACTGGTACA CTAAAGAG AATTTGGACA
12001 TGATGGACA RATTAGTTA ATTGTTTTT AAAGTTTGA CCACCTTAC ATATCTGAA TAAATCTAC TTGTTGTGG TGTATATTT TTGATACATT
ACTACCTGT TTAATGCAAT TAACAAAAG TTGCAACT GTGGGAATG TAPAGACCTT AATTAAGATG ACCAACACC ACATATAAAA AACTATGTAA
12101 CTGGATCT TTTGCTAAT ATTTGTTGA AATGTTTGT AICTTTGTT ATGAGAGATA TTGGTCTGT GTTTCTCTTT CTGTATATGT CATTTCTAG

Figure 7-7

GATCCTAAGA AATCGGATTA TAAACAACT TTTACAACA TAGAACAAG TACTCTCTAT AACAGACAA CAAAGGAAA GACATTACA GTAAAGATC
12201 TTCCGGTATT AAGGTAATGC TGGCCTAGTT GAATGATTTA GGAAGTATTC CCTCTGCTTC TGTCTTCTGA AGCGGAAGAG CGCCCAATAC GCAAAACGGCC
AAGGCCATAA TTCCATTAG ACCGGATCAA CTTACTAAT CTTTATAG GGAGACGAAG ACAGAGACT TCGCCTTCTC GCGGGTTATG CGTTTGGGGG
12301 TCTCCCGCG CGTTGGCCGA TTCAATATG CAGCTGGCAC GACAGTTTC CCGACTGGAA AGCGGCGAGT GAGCGCAACG CAATTAATGT GAGTTAGCTC
AGAGGGGCGC GCACCGGCT AAGTAATTAC GTCGACCGTG CTGTCCAAG GGTGACCTT TCGCCCGTCA CTCGCGTGC GTTAATTACA CTCAAATCGAG
12401 ACTCATTAGG CACCCAGGC TTACACTTT ATGCTTCCGG CTCGTATGTT GTGTGCAATT GTGAGGGAT AACAAATTC CACAGGAAC AGCTATGACA
TGAGTAATCC GTGGGTCGG AATGTGAAA TACGAGGCG GAGCATACAA CACACCTTAA CACTCGCTA TTGTTAAGT GTGTCTTTC TCGATACTGT
12501 TGATTACGAA TTAA
ACTAATGCTT AATT

>length: 12514

Figure 7-8

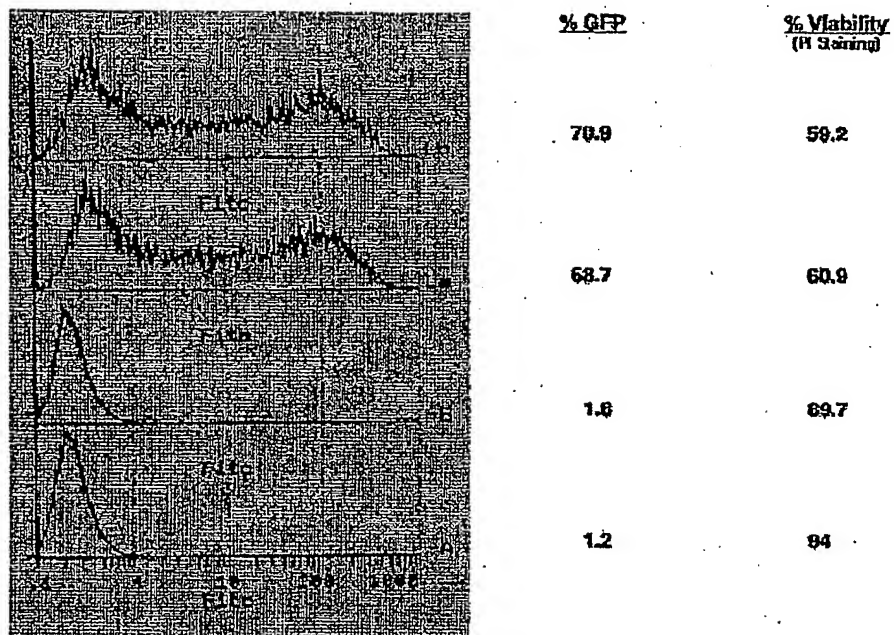


Figure 8. FACS analysis of transiently transfected CHO cells with a GFP plasmid in 250 ml spinner transfection.

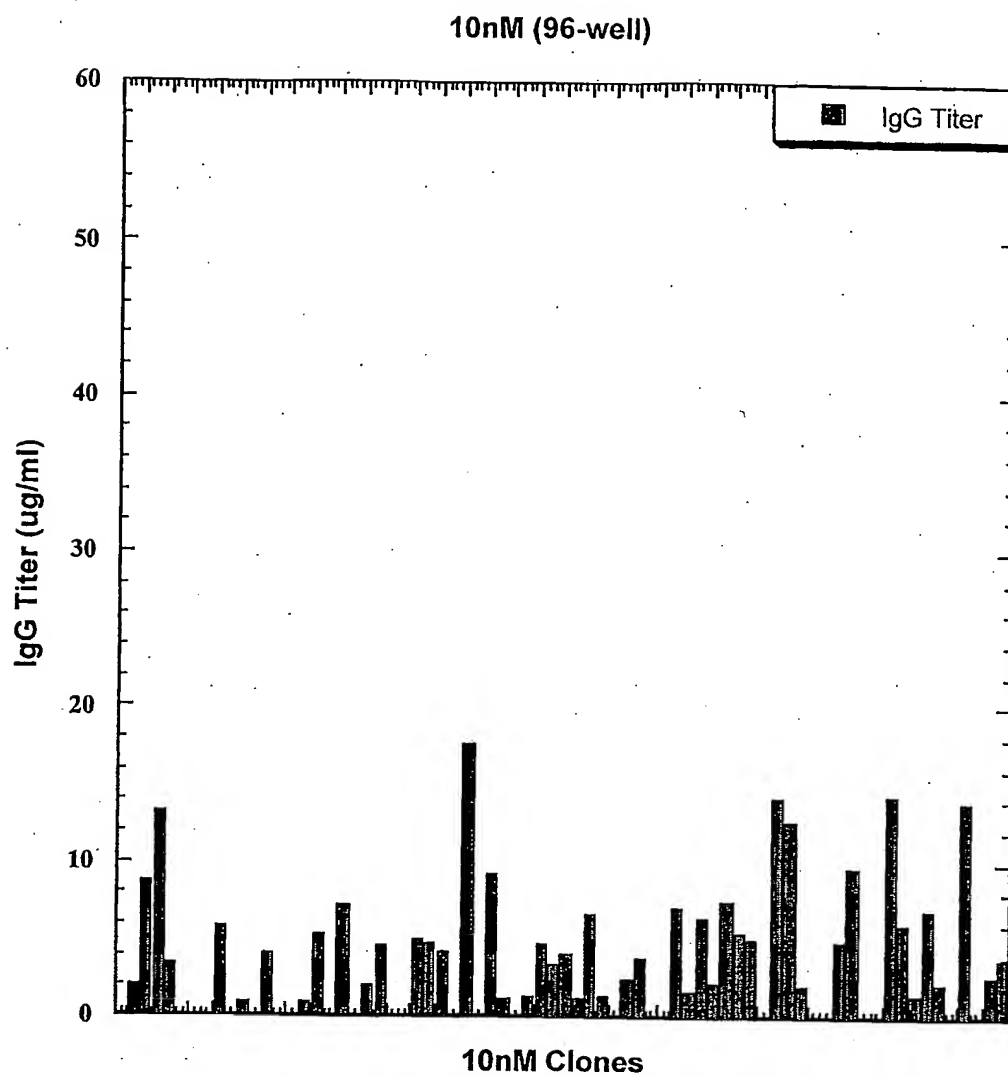


Figure 9. Expression level of clones from traditional 10 nM MTX selection.

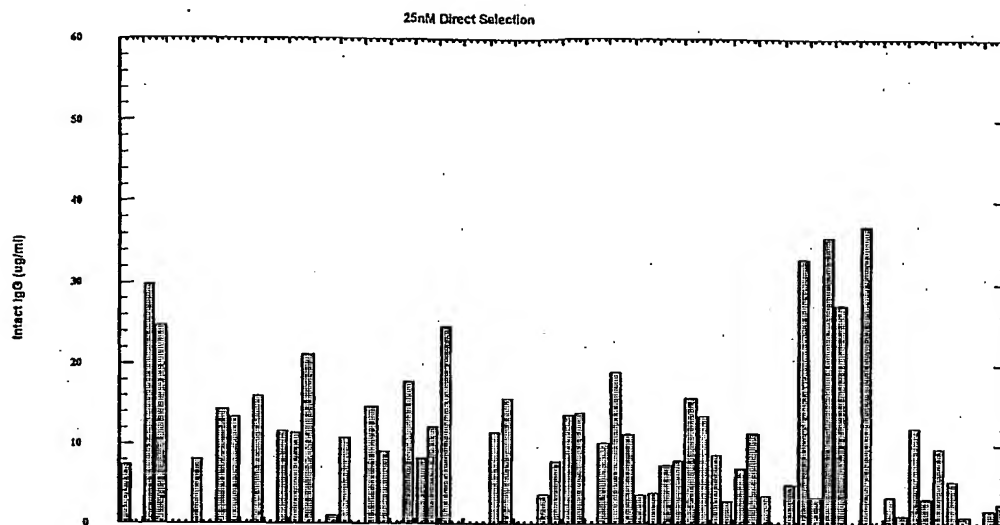


Figure 10-1

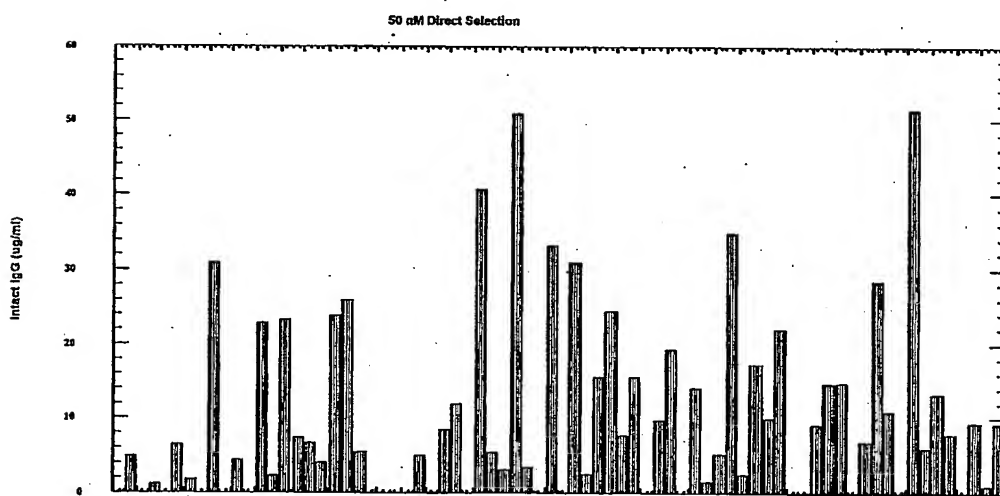


Figure 10-2

Figures 10.1 and 10.2. Expression level of clones from 25 and 50 nM MTX direct selections of SV40-based constructs derived from spinner transfection, respectively.

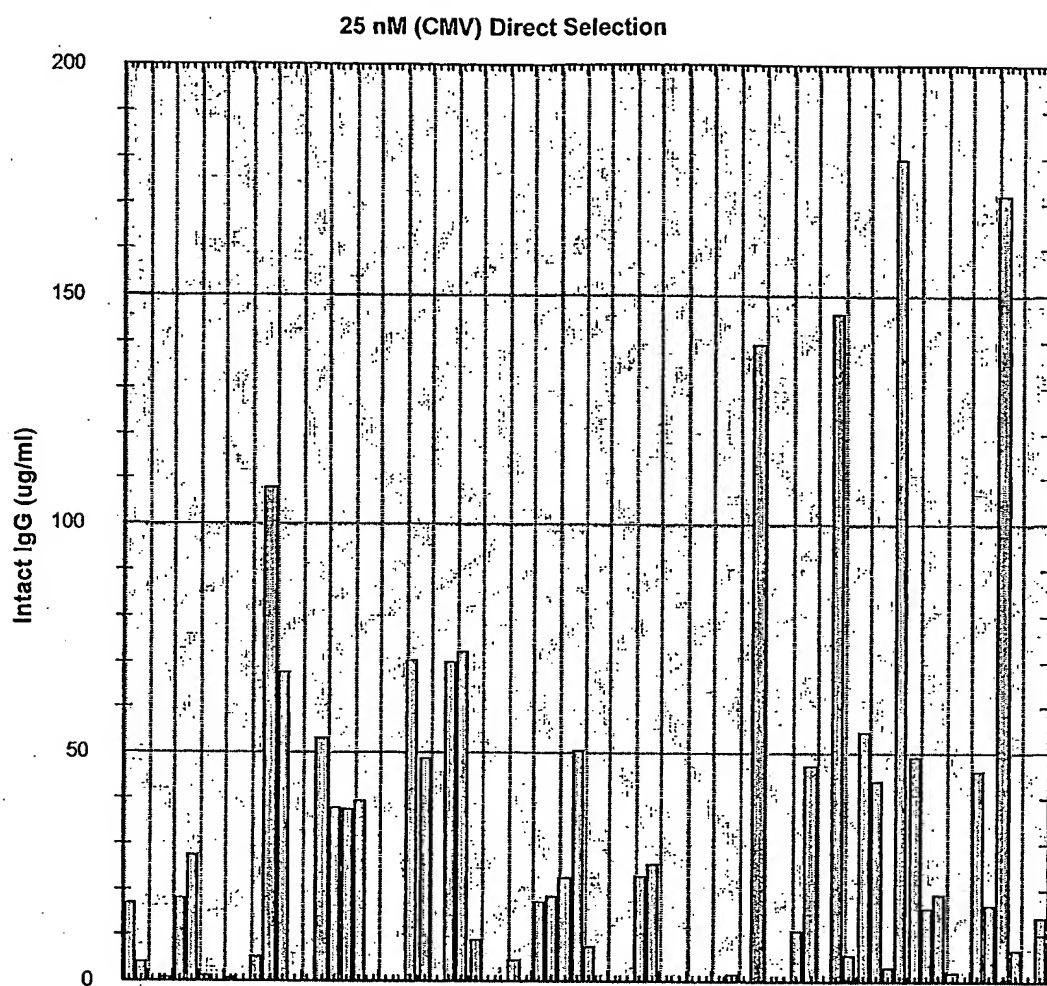


Figure 11. Expression level of clones from 25 nM MTX direct selection of CMV construct derived from spinner transfection.

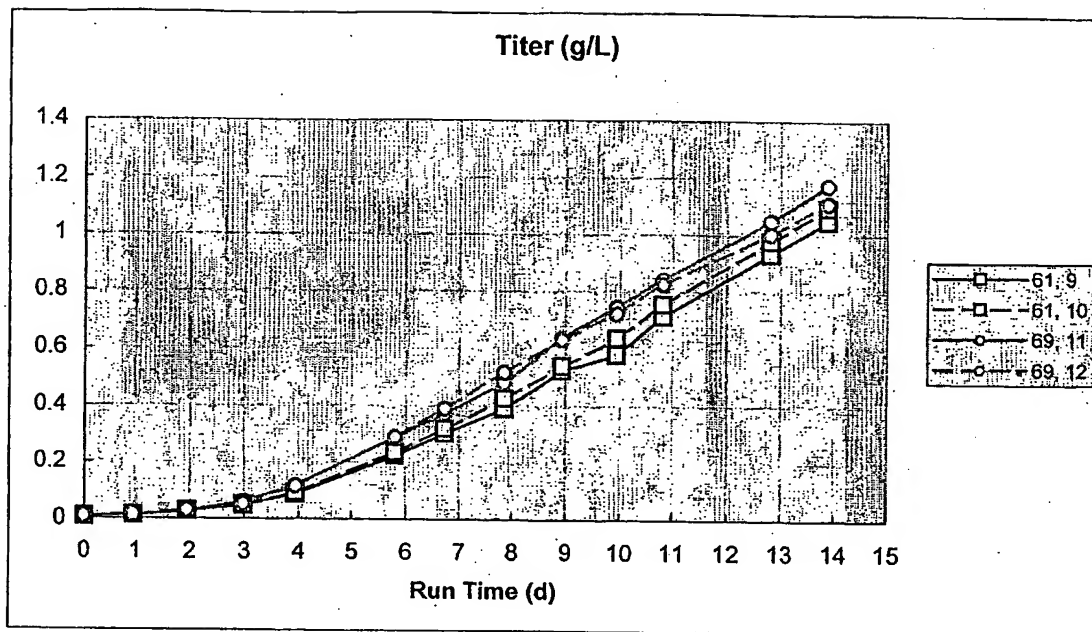


Figure 12. Titer Evaluation in Miniferm.

Figure 13. Plasmid pCMV.IPD.Heterologous Polypeptide

5 <400>
60 TTCGAGCTCG CCCGACATTG ATTATTGACT AGAGTCGATC ACCGGTAGTA ATCAATTACG
120 GGGTCAATTAG TTCATAGCCC ATATATGGAG TTCCGCGTTA CATAACTTAC GGTAAATGGC
180 CCGCTGGCT GACCGCCCAA CGACCCCGC CCATTGACGT CAATAATGAC GTATGTTCCC
240 ATAGTAACGC CAATAGGAC TTTCATTGA CGTCAATGGG TGGAGTATTT ACGGTAAACT
300 GCCCCACTGG CAGTACATCA AGTGTATCAT ATGCCAAGTA CGCCCCCTAT TGACGTCAAT
360 GACGGTAAAT GGCCCGCCTG GCATTATGCC CAGTACATGA CCTATGGGA CTTTCCTACT
420 TGGCAGTACA TCTACGTATT AGTCATCGCT ATTACCATGG TGATGCGGTT TTGGCAGTAC
480 ATCAATGGGC GTGGATAGCG GTTGACTCA CGGGGATTTC CAAGTCTCCA CCCATTGAC
540 GTCAATGGGA GTTGTGTTTG GCACCAAAAT CAACGGGACT TTCCAAAATG TCGTAACAAC
600 TCCGCCCCAT TGACGCAAT GGGCGGTAGG CGGTACGGT GGGAGGTCTA TATAAGCAGA
660 GCTCGTTTAG TGAACCGTCA GATCGCCTGG AGACGCCATC CACGCTGTTT TGACCTGGGC
720 CCGGCCGAGG CCGCCTCGGC CTCTGAGCTA TTCCAGAAGT AGTGAGGAGG CTTTTTTGGA
780 GGCCTAGGCT TTTGCAAAA GCTAGCTTAT CCGGCCGGGA ACGGTGCATT GGAACGCGGA
840 TTCCCCGTGC CAAGAGTGAC GTAAGTACCG CCTATAGAGC GACTAGTCCA CCATGACCGA
900 GTACAAGCCC ACGGTGCGCC TCGCCACCCG CGACGACGTC CCGGGGGCGG TAGGCACCCCT

Figure 13.1

960 CGCCGCGCGG TTGCGCGACT ACCCGGCAC GCGCCACACC GTAGACCCGG ACCGCCACAT
1020 CGAGCGGGTC ACCGAGCTGC AAGAACTCTT CCTCACGGCG GTCGGGCTCG ACATCGGCAA
1080 GGTGTGGGTC GCGGACGACG GCGCGGCGGT GCGGCTCTGG ACCACGCCGG AGAGCGTCGA
1140 AGCGGGGGCG GTGTTGCGG AGATCGGCGC GCGCATGGCC GAGTTGAGCG GTTCCCGGCT
1200 GGCGCGCAG CAACAGATGG AAGGCCTCCT GCGCGCGCAC CGGCCCAAG AGCCCGCGTG
1260 GTTCTGGCC ACCGTCGGG TCTCGCCCGA CCACCAGGC AAGGCTCTGG GCAGCGCCGT
1320 CGTGCTCCCC GGAGTGGAGG CGCGCGAGCG CGCCGGGGTG CCCGCCCTCC TGGAGACCTC
1380 CGCGCCCCGC AACCTCCCT TCTACGAGG GCTCGGCTTC ACCGTCACCG CCGACGTCGA
1440 GGTGCGCGAA GGACCGCGCA CCTGGTGCAT GACCGGCAAG CCCGGTGCCA ACATGGTTCG
1500 ACCATTGAAC TGATCGTCG CCGTGTCOCA AAATATGGG ATTGGCAAGA ACGGAGACCT
1560 ACCCTGGCCT CCGCTCAGGA ACGCGTTCAA GTACTTCCAA AGAATGACCA CAACCTCTTC
1620 AGTGAAGGT AAACAGAATC TGGTGATTAT GGGTAGGAAA ACCTGGTTCT CCATTCTGA
1680 GAAGATCGA CCTTTAAAGG ACAGATTAA TATAGTTCTC AGTAGAGAC TCAAAGAACC
1740 ACCACGAGGA GCTCATTTTC TTGCCAAAAG TTTGGATGAT GCCTTAAGAC TTATTGAACA
1800 ACCGGAATTG GCAAGTAAAG TAGACATGGT TTGGATAGTC GGAGGCAGTT CTGTTTACCA
1860 GGAAGCCATG AATCAACCAG GCCACCTCAG ACTCTTTGTG ACAAGGATCA TGCAGGAATT
1920 TGAAGTGCAC ACGTTTTTCC CAGAAATTGA TTTGGGGAAA TATAACCTC TCCCAGATA
1980 CCCAGCGTC CTCTCTGAGG TCCAGGAGGA AAAAGGCATC AAGTATAAGT TTGAAGTCTA

Figure 13.2

2040 CGAGAAGAAA GACTAACGTT AACTGCTCCC CTCCTAAAGC TATGCATTTT TATAAGACCA
2100 TGAGACTTTT GCTGGCTTTA GATCCCTTG GCTTCGTTAG AACGCAGCTA CAATTAATAC
2160 ATAACCTTAT GTATCATACA CATACGATTT AGGTGACACT ATAGAATAAC ATCCACTTTG
2220 CCTTCTCTC CACAGGTGTC CACTCCAGG TCCACTGCA CCTCGGTTCT ATCGATTGAA
TTCACC --Insert Sequence of Interest--
CGA TGGCGGCCAT GGCCCAACTT GTTATTGCA GCTTATAATG
GTTACAATA AAGCAATAGC ATCACAATT TCACAAATAA AGCATTTTT TCACTGCATT
CTAGTTGGG TTTGTCCAA CTCATCAATG TATCTTATCA TGCTGGATC GGAATTAAT
TCGGGCAGC ACCATGGCCT GAAATAACCT CTGAAAGAGG AACTTGGTTA GGTACCTATT
AATAGTATC AATTACGGG TCATTAGTTC ATAGCCATA TATGGAGTTC CGGTTACAT
AACTTACGGT AAATGGCCG CCTGGCTGAC CGCCCAACGA CCCCCGCCA TTGACGTCAA
TAATGACGTA TGTTCCATA GTAAAGCCAA TAGGGACTTT CCATTGACGT CAATGGGTGG
AGTATTTACG GTAAACTGCC CACTTGGCAG TACATCAAGT GTATCATATG CCAAGTACGC
CCCCATTTGA CGTCAATGAC GGTAAATGGC CCGCCTGGCA TTATGCCAG TACATGACCT
TATGGGACTT TCCTACTGG CAGTACATCT ACGTATTAGT CATCGCTATT ACCATGGTGA
TGCGGTTTG GCAGTACATC AATGGGCGTG GATAGCGGTT TGACTCACGG GGATTTCCAA
GTCTCCACCC CATTGACGTC AATGGGAGTT TGTTTTGGCA CCAAAATCAA CGGGACTTTC
CAAAATGTCG TAACAACCTCC GCCCATTTGA CGAAATGGG CGGTAGGCGT GTACGGTGGG

Figure 13.3

AGGTCTATAT AAGCAGAGCT CGTTTAGTGA ACCGTCAGAT CGCCTGGAGA CGCCATCCAC
GCTGTTTTGA CTTGCTAGCT TATCGGGCCG GGAACGGTGC ATTGGAACGC GGATTCCCCG
TGCCAAGAGT CAGGTAAGTA CCGCCTATAG AGTCTATAGG CCCACCCCTT TGGCTTCGTT
AGAACGGGC TACAATTAAT ACATAACCTT TTGGATCGAT CCTACTGACA CTGACATCCA
CTTTTCTTT TTCTCCACAG GTGTCCACTC CCAGGTCCAA CTGCACCTCG GTTCGGGAAG
CTCGCTGGG CTGCATCGAT TGAATTCCAC C --Insert Sequence of Interest--
CGATGG CCGCCATGGC CCAACTTGT TATTGCAGCT TATAATGGTT
ACAAATAAAG CAATAGCATC ACAAATTCA CAAATAAAGC ATTTTTCCTA CTGCATTCTA
GTTGTGGTT GTCCAACTC ATCAATGTAT CTTATCATGT CTGGATCGGG AATTAATTCG
GGCAGCACCC ATGGCCTGAA ATAGTTTAA ACCCTCTGAA AGAGGAACTT GGTTAGGTAC
CGACTAGTCT TTTGCAAAAA GCTGTTACCT CGAGCGGCGG CTTAATTAAAG GCGCGCCATT
TAAATCCTGC AGGTACAGC TTGGCACTGG CCGTCGTTTT ACAACGTCGT GACTGGGAAA
ACCCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC CCCTTTGCGC AGCTGGCGTA
ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT
GGCGCCTGAT GCGGTATTT CTCCTTACGC ATCTGTGCGG TATTTCACAC CGCATACGTC
AAAGCAACCA TAGTACGCGC CCTGTAGCGG CGCATTAAAG CCGGCGGGTG TGGTGGTTAC
GGCAGCGTG ACCGCTACAC TTGCCAGCGC CCTAGCGCCC GCTCCTTTTCG CTTTCTTCCC
TTCTTTCTC GCCACGTTG CCGGCTTTCC CCGTCAAGCT CTAAATCGGG GGCTCCCTTT

Figure 13.4

AGGGTTCCGA TTTAGTGCTT TACGGCACCT CGACCCCAA AACTTGATT TGGTGATGG
TTACAGTAGT GGGCCATCGC CCTGATAGAC GGTTTTTCGC CCTTTGACGT TGGAGTCCAC
GTTCTTTAAT AGTGGACTCT GTTCCAAAC TGAACAACA CTCAAACCCTA TCTCGGGCTA
TTCCTTTGAT TTATAAGGA TTTTGCCGAT TTCGGCCTAT TGGTTAAAA ATGAGCTGAT
TTAACAAAA TTTAACGCGA ATTTTAACAA AATATTACG TTTACAAATT TATGGTGCAC
TCTCAGTACA ATCTGCTCTG ATGCCGCATA GTTAAGCCAG CCCCGACACC GCCCGGACAC
CGGCCAACAC CCGCTGACGC GCCCTGACGG GCTTGCTCTGC TCCCGGCATC CGCTTACAGA
CAAGCTGTGA CCGTCTCCGG GAGCTGCATG TGTACAGAGT TTTACCCGTC ATCACCAGAA
CGCGGAGAG ACGAAGGGC CTCGTGTATC GCCTATTTT ATAGGTTAAT GTCATGATAA
TAATGGTTTC TTAGACGTCA GGTGGCACTT TTCGGGAAA TGTGCGCGGA ACCCCTATTT
GTTTATTTT CTAATACAT TCAATATGT ATCCGCTCAT GAGACATAA CCTGATAAA
TGCTTCAATA ATATTGAAA AGGACAGTA TGAGTATTCA ACATTTCCGT GTCGCCCTTA
TTCCTTTT TGCGGCATTT TGCCCTCCTG TTTTGTCTCA CCCAGAAACG CTGGTAAAG
TAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA
GCGTAAGAT CCTTGAGAGT TTTGCCCCG AAGAACGTTT TCCAATGATG AGCCTTTTA
AAGTTCTGCT ATGTGGCGG GTATTATCCC GTATTGACGC CGGCAAGAG CAATCGGTC
GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCACTCACA GAAAGCATC
TTACGGATGG CATGACAGTA AGAATATT GCAGTCTGTC CATACCATG AGTGATAACA

Figure 13.5

CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC
ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA
TACCAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAAC
TATTAATCTGG CGAACTACTT ACTCTAGCTT CCCGGGCAACA ATTAATAGAC TGGATGGAGG
CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG
ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG
GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC
GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC
AAGTTTACTC ATATAACTT TAGATTGATT TAAACTTCA TTTTAAATT AAAAGGATCT
AGGTGAAGAT CCTTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC
ACTGAGGCTC AGACCCCGTA GAAAAGATCA AAGGATCTC TTGAGATCCT TTTTTCCTGC
GCGTAATCTG CTGCTTGCAA ACAAAAAAC CACCGGTACC AGCGGTGGT TGTTTGCCGG
ATCAAGAGCT ACCAATCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA
ATACTGTTCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC
CTACATACT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT
GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATAA GGGCAGCGG TCGGGCTGAA
CGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC
TACAGCGTGA GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC

Figure 13.6

CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAAACGCCT
GGTA^uCTTTA TAGTCTGTG GGGTTTCGCC ACCTCTGACT TGAGCGTCTGA TTTTGTGAT
GCTCGTCAGG GGGCGGAGC CTATGGAAAA AGCCAGCAA CGCGGCCTTT TTACGGTTCC
TGGCCTTTTG CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCT GATTCTGTGG
ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC
GCAGCGAGTC AGTGAGCGAG GAAGCGGAG AGCGCCCAAT ACGCAAACCG CCTCTCCCCG
CGCGTTGGCC GATTCAITAA TGCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGGCA
GTGAGCGCAA CGCAATTAA GTGAGTTAGC TCACTCATTA GGCACCCAG GCTTTACT
TTATGCTTCC GGCTCGTAG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA
ACAGCTATGA CATGATTACG AATTAA

Figure 13.7

Figure 14. Plasmid SV40.IPD.Heterologous Polypeptide

6 <400>
60 TTCGAGCTCG CCGACATTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT
120 CAGTTAGGGT GTGGAAAGTC CCCAGGCTCC CCAGCAGGCA GAAGTATGCA AAGCATGCAT
180 CTCATTAGT CAGCAACCAG GTGTGGAAG TCCCCAGGCT CCCAGCAGG CAGAAGTATG
240 CAAAGCATGC ATCTCAATTA GTCAGCAACC ATAGTCCCGC CCCTAACTCC GCCCATCCCG
300 CCCCTAACTC CGCCAGTTC CGCCCATTTCT CCGCCCCATG GCTGACTAAT TTTTTTATT
360 TATGCACAGG CCGAGGCCGC CTCGGCCCTCT GAGCTATTCC AGAAGTAGTG AGGAGGCTTT
420 TTTGGAGGCC TAGGCTTTTG CAAAAGCTA GCTTATCCGG CCGGGAACGG TGCATTGGAA
480 CGCGGATTCC CCGTGCCAAG AGTGACGTAA GTACCCGCTA TAGAGCGACT AGTCCACCAT
540 GACCGAGTAC AAGCCACGG TGCGCCTCGC CACCCGCGAC GACGTCCCGC GGGCCGTACG
600 CACCCTCGCC GCCGCGTTG CCGACTACCC CGCCACGCGC CACACCGTAG ACCCGGACCG
660 CCACATCGAG CGGGTCACCG AGCTGCAAGA ACTCTTCTC ACGCGCGTCG GGCTCGACAT
720 CGGCAAGGTG TGGGTCGCGG ACGACGCGC CGCGGTGGCG GTCTGGACCA CGCCGGAGAG
780 CGTCGAAGCG GGGCGGTGT TCGCCGAGAT CGGCCCGCGC ATGGCCGAGT TGAGCGGTTT
840 CCGCTGGCC GCGCAGCAAC AGATGGAAG CCTCCTGGCG CCGCACCGGC CCAAGAGGCC
900 CGCGTGGTTC CTGGCCACCG TCGGCGTCTC GCCCGACCA CAGGGCAAGG GTCTGGGCAG

Figure 14.1

960 CGCCGTCGTG CTCCCCGGAG TGGAGGGCGC CGAGCGGCGC GGGGTGCCCC CTTCTCTGGA
1020 GACCTCCGGC CCCCAGCAAC TCCCCCTTCTA CGAGCGGCTC GGCTTCAACG TCACCGCCGA
1080 CGTCGAGTGC CCGAAGGACC GCGGACCTG GTGCATGACC CGCAAGCCCC GTGCCAACAT
1140 GGTTCGACCA TTGAACTGCA TCGTCGCGGT GTCCCAAAAT ATGGGGATTG GCAAGAACGG
1200 AGACCTACCC TGCCCTCCGC TCAGGAACGC GTTCAAGTAC TTCCAAAGAA TGACCACAAC
1260 CTCTTCAGTG GAAGGTAAAC AGAATCTGGT GATTATGGGT AGGAAAACCT GGTCTCTCCAT
1320 TCCTGAGAAG AATCGACCTT TAAAGGACAG AATTAATATA GTTCTCAGTA GAGAACTCAA
1380 AGAACCACCA CGAGGAGCTC ATTTTCTTGC CAAAAGTTTG GATGATGCCT TAAGACTTAT
1440 TGAACAACCG GAATTGGCAA GTAAAGTAGA CATGGTTTGG ATAGTCGGAG GCAGTTCTGT
1500 TTACCAGGAA GCCATGAATC AACCAGGCCA CCTTAGACTC TTGTGACAA GGATCATGCA
1560 GGAATTGTGA AGTGACACGT TTTTCCCAGA AATTGATTG GGGAAATATA AACCTCTCCC
1620 AGAATACCCA GCGTCCCTCT CTGAGGTCCA GGAGGAAAAA GGCATCAAGT ATAAGTTTGA
1680 AGTCTACGAG AAGAAAGACT AACGTTAACT GCTCCCCCTCC TAAAGCTATG CATTTTATA
1740 AGACCATGGG ACTTTTGCTG GCTTTAGATC CCGTTGGCTT CGTTAGAACG CAGCTACAAT
1800 TAATACATAA CCTTATGTAT CATAACATA CGATTAGGT GACACTATAG ATAACATCCA
1860 CTTTGCCCTT CTCTCCACAG GTGTCCACTC CCAGGTCCAA CTGCACCTCG GTTCTATCGA
1920 TTGAATTCCA CC -Insert Sequence of Interest-
CGATGGCC GCCATGGCCC AACTTGTTTA TTGCAGCTTA

Figure 14.2

TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTTCACA AATAAAGCAT TTTTTCAC
GCATTCTAGT TGTGGTTTGT CCAAACTCAT CAATGTATCT TATCATGTCT GGATCGGAA
TTAATTCGGC GCAGCACCAT GGCTGAAAT AACCTCTGAA AGAGGAACTT GGTAGGTAC
CTCTGAGGC GGAAGAACC AGCTGTGGAA TGTGTGTCAG TTAGGGTGTG GAAAGTCCCC
AGGCTCCCCA GCAGGCAGAA GTATGCAAAG CATGCATCTC AATTAGTCAG CAACCAGGTG
TGGAAAGTCC CCAGGCTCCC CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC
AGCAACCATTA GTCCCGCCCC TAACTCCGCC CATCCGGCC CTAATCCGC CCAGTCCGC
CCATTCTCG CCCCATGGCT GACTAATTTT TTTTATTAT GCAGAGGCCG AGGCCGCCCTC
GGCTCTGAG CTATTCAGA AGTAGTGAGG AGGCTTTTTT GGAGAGCTT TTGCAAAAAG
CTAGCTTATC CGGCCGGAA CGGTGCATTG GAACGGGAT TCCCGTGCC AAGAGTCAGG
TAAGTACCGC CTATAGAGTC TATAGGCCCA CCCCCTTGGC TTCGTTAGAA CGCGGCTACA
ATTAATACAT AACCTTTTGG ATCGATCCTA CTGACACTGA CATCCACTTT TTCCTTTTCT
CCACAGGTGT CCACTCCCAG GTCCAACTGC ACCTCGGTTT CCGAAGCTAG CTTGGGCTGC
ATCGATTGAA TTCCACC -Insert Sequence of Interest-
CGATGGCCGC CATGGCCCAA CTTGTTTATT GCAGTTATA ATGGTTACAA ATAAAGCAAT
AGCATCACAA ATTTACAAA TAAGCATTT TTTTCACTGC ATTCTAGTTG TGGTTGTCC
AAACTCATCA ATGTATCTTA TCACTCTGAG ATCGGGAATT AATTCGGCGC AGCACCATGG
CCTGAAATAA GTTTAAACCC TCTGAAAGAG GAACCTGGTT AGGTACCGAC TAGTCTTTTG

Figure 14.3

CAAAAGCTG TTACCTCGAG CGGCGGCTTA ATTAAGGCGC GCCATTAAAT TCCTGCAGGT
AACAGCTTGG CACTGGCCGT CGTTTTACAA CGTCGTGACT GGGAAAACCC TGGCGTTACC
CAACTTAATC GCCTTGCAGC ACATCCCCCT TTCGCCAGCT GCGGTAATAG CGAAGAGGCC
CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTGAATG GCGAATGGCG CCTGATGCGG
TATTTTCTCC TTACGCATCT GTGCGGTATT TCACACCGCA TACGTCAAAG CAACCATAGT
ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT GGTACGCGC AGCGTGACCG
CTACACTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT CTTCCTTCC TTTCTCGCCA
CGTTGCGCGG CTTTCCCGGT CAAGCTCTAA ATCGGGGGCT CCCTTTAGGG TTCCGATTTA
GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTGGG TGATGGTTCA CGTAGTGGG
CATCGCCCTG ATAGACGGTT TTTCCGCCCT TGACGTTGGA GTCCACGTTT TTTAATAGTG
GACTCTTGT CCAAACTGGA ACAACACTCA ACCCTATCTC GGGCTATTCT TTTGATTAT
AAGGGATTT GCCGATTTCG GCCTATTGGT TAAAAAATGA GCTGATTAA CAAAAATTA
ACGCGAATT TAACAAAATA TTAACGTTTA CAATTTTATG GTGCACTCTC AGTACAATCT
GCTCTGATGC CGCATAGTTA AGCCAGCCCC GACACCCGCC AACACCCGCT GACGCGCCCT
GACGGGCTG TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT
GCATGTCTCA GAGGTTTCA CCGTCATCAC CGAAACGGCG GACGAAAGGG CCTCGTGATA
CGCCTATTTT TATAGGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT
TTTCGGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG

TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT
ATGAGTATTC AACATTTCCG TGTGCCCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT
GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAGATG CTGAAGATCA GTTGGGTGCA
CGAGTGGGTT ACATCGAAT GGATCTCAAC AGCGTAAGA TCCTTGAGAG TTTTCGCCCC
GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC
CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG
GTTGAGTACT CACCAGTCAC AGAAAGCAT CTACGGATG GCATGACAGT AAGAGAATTA
TGCAGTGTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT
GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG
CCTGTAGCAA TGGCAACAAC GTTGCGCNA A CTATTAACTG GCGAACTACT TACTCTAGCT
TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC
TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT
CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC
ACGACGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC
TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT
TTAAACTTC ATTTTAAAT TAAAAGGATC TAGGTGAAGA TCCTTTTTGA TAATCTCATG
ACCAAAATCC CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGATC

Figure 14.5

AAAGGATCTT CTTGAGATCC TTTTTCCTG CGCGTAATCT GCTGCTGCA AACAAAAAA
CCACCGCTAC CAGCGGTGGT TTGTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG
GTAAGTGGCT TCAGCAGAGC GCAGATACCA AATAGTGCC TTCTAGTGTG GCCGTAGTTA
GGCCACCACT TCAAGAACTC TGTAGCACCG CTTACATACC TCGCTCTGCT AATCCTGTTA
CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGCTTTACCG GGTGGACTC AAGACGATAG
TTACCGGATA AGCGCGAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTTG
GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG ACCTATGAGA AAGCGCCACG
CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG
CGCAGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT CGGGTTTCGC
CACCTCTGAC TTGAGCGTCG ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA
AACGCCAGCA ACGCGGCCTT TTTACGGTTC CTGGCCCTTT GCTGGCCCTT TGCTCACATG
TTCTTTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT
GATACCGCTC GCCGAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA
GAGCGCCAA TACGCAAAACC GCCTCTCCCC GCGCGTTGGC CGATTCAATA ATGCAGCTGG
CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAG
CTCACTCATT AGGCACCCCA GGCTTTACAC TTTAIGTTC CGGCTCGTAT GTTGTGTGGA
ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACATGATTAC GAATTAA

Figure 14.6

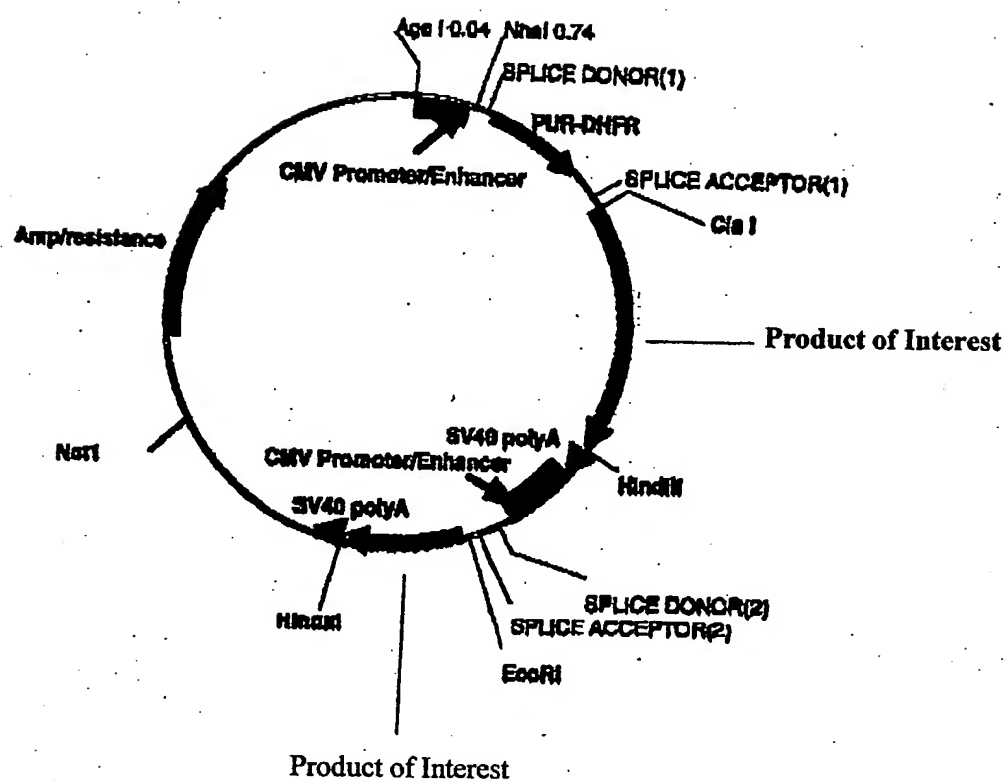


Figure 15. pCMV.IPD.HP

Timeline and Titer Comparison

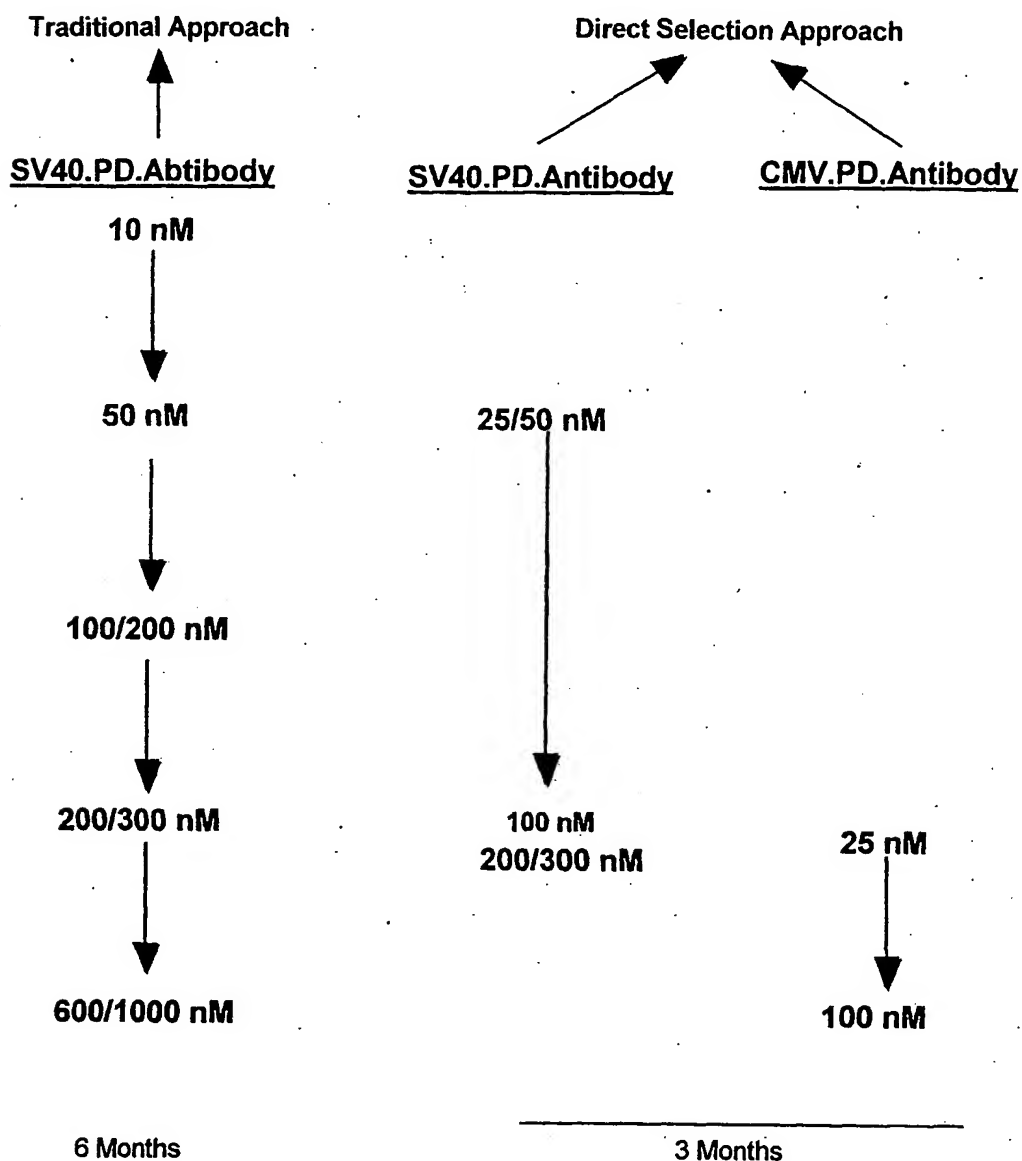


Figure 16. Timeline and Titer Comparison.

SEQUENCE LISTING

<110> Krummen, Lynne
Shen, Amy
Chisum, Venessa

<120> INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-
EXPRESSING
PRODUCTION CELL LINES

<130> 22338/00101

<150> US 60/426,095

<151> 2002-11-14

<160> 4

<170> PatentIn version 3.1

<210> 1

<211> 5147

<212> DNA

<213> Artificial

<220>

<223> plasmid pSV.IPUR circular ds-DNA

<400> 1

ttcgagctcg cccgacattg attattgact agagtcgacgac gacagctgtg gaatgtgtgt	60
cagttagggt gtggaaagtc ccagagctcc ccagcaggca gaagtatgca aagcatgcat	120
ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg	180
caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gccatcccg	240
cccctaactc cgcccagttc cgcccattct ccgcccctg gctgactaat tttttttatt	300
tatgcagagg ccgaggccgc ctggcctct gagctattcc agaagtagtg aggaggcttt	360
tttgagggcc taggcttttg caaaaagcta gcttatccgg ccgggaacgg tgcattggaa	420
cgcggtattcc ccgtgccaaag agtgacgtaa gtaccgccta tagagcgact agtccacat	480
gaccgagtag aagcccacgg tgcgcctcgc caccgcgcgac gacgtcccc gggccgtacg	540
caccctcgcc gccgcgttcg ccgactaccc cgccacgcgc cacaccgtcg acccgagccg	600
ccacatcgag cgggtcaccg agctgcaaga actcttctc acgcgcgtcg ggctcgacat	660
cggaaggtg tgggtcgcg acgacggcgc cgcggtggcg gtctggacca cgccgagag	720
cgctgaagcg ggggcggtgt tcgccgagat cggccgcgc atggccgagt tgagcggttc	780
ccggtggcc gcgcagcaac agatggaagg cctctggcg ccgcaccggc ccaaggagcc	840
cgcggtgttc ctggccaccg tcggcgtctc gcccgaccac cagggaagg gtctgggcag	900

cgccgtcgtg ctccccggag tggaggcggc cgagcgcgcc ggggtgccc ccttcctgga 960
gacctccgcg ccccgcaacc tccccctcta cgagcggctc ggcttcaccg tcaccgccga 1020
cgtcgagtgc ccgaaggacc gcgcgacctg gtgcatgacc cgcaagcccg gtgcctgagt 1080
taactgctcc cctcctaaag ctatgcattt ttataagacc atgggacttt tgctggcttt 1140
agatccccctt ggcttcgtta gaacgcagct acaattaata cataacctta tgtatcatac 1200
acatacgatt taggtgacac tatagataac atccactttg cttttctctc cacagggtgc 1260
cactcccagg tccaactgca cctcggttct atcgattgaa ttccccgggg atcctctaga 1320
gtcgacctgc agaagcttcg atggccgcca tggcccaact tgtttattgc agcttataat 1380
ggttacaaat aaagcaatag catcacaat ttcacaaata aagcattttt ttcactgcat 1440
tctagtgtg gtttgtcaa actcatcaat gtatcttctc atgtctggat cgatcgggaa 1500
ttaattcggc gcagcaccat ggcctgaaat aacctctgaa agaggaaactt ggtaggtac 1560
cttctgaggc ggaaagaacc agctgtggaa tgtgtgtcag ttaggggtgtg gaaagtcccc 1620
aggctcccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag caaccagggtg 1680
tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcatc tcaattagtc 1740
agcaaccata gtcccgcccc taactccgcc catcccgccc ctaactccgc ccagttccgc 1800
ccattctccg ccccatggct gactaatttt ttttatttat gcagaggccg aggcgcctc 1860
ggcctctgag ctattccaga agtagtgagg aggccttttt ggaggcctag gcttttgcaa 1920
aaagctgtta cctcgagcgg ccgcttaatt aaggcgcgcc atttaaacc tgcaggtaac 1980
agcttggcac tggccgtcgt tttaaacgt cgtgactggg aaaaccctgg cgttacccaa 2040
cttaatcgcc ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggcccgc 2100
accgatcgcc cttccaaca gttgcgtagc ctgaatggcg aatggcgct gatgcggtat 2160
tttctcctta cgcactgtg cggtatttca caccgcatac gtcaaagcaa ccatagtacg 2220
cgccctgtag cggcgcatca agcgcggcgg gtgtgggtgt tacgcgcagc gtgaccgcta 2280
cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt ctgcgccagt 2340
tcgcgggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg 2400
ctttaoggca cctcgacccc aaaaaacttg atttgggtga tggttcacgt agtgggccat 2460
cgccctgata gacggttttt cgccctttga cgttggagtc cacgttcttt aatagtggaac 2520
tcttgttcca aactggaaca aactcaacc ctatctcggg ctattctttt gatttataag 2580

ggattttgcc gatttcggcc tatttggttaa aaaatgagct gatttaacaa aaatttaacg 2640
cgaatttttaa caaaatatta acgtttacaa ttttatggtg cactctcagt acaatctgct 2700
ctgatgccgc atagttaagc caactccgct atcgctacgt gactgggtca tggctgcgcc 2760
ccgacacccg ccaacacccg ctgacgcgcc ctgacgggct tgtctgctcc cggcatccgc 2820
ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt cagaggtttt caccgtcatc 2880
accgaaacgc gcgaggcagt attcttgaag acgaaagggc ctcgtgatac gcctatTTTT 2940
ataggTTaat gtcataataa taatggtttc ttagacgtca ggtggcactt ttcggggaaa 3000
tgtgcgcgga acccctatTTt gtttatTTTT ctaaatacat tcaaatatgt atccgctcat 3060
gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca 3120
acatttcgct gtcgccctta ttcctTTTTt tgcggcattt tgccttcctg tttttgctca 3180
cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta 3240
catcgaactg gatctcaaca gcggttaagat ccttgagagt tttcgccccg aagaacgTTt 3300
tccaatgatg agcactTTTT aagttctgct atgtggcgcg gtattatccc gtgatgacgc 3360
cgggcaagag caactcggtc gcgcataca ctattctcag aatgacttgg ttgagtactc 3420
accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgtctc 3480
cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa 3540
ggagctaacc gctTTTTtgc acaacatggg ggatcatgta actcgccttg atcgttggga 3600
accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc cagcagcaat 3660
ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt cccggcaaca 3720
attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc 3780
ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc gcggtatcat 3840
tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag 3900
tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggTgcct cactgattaa 3960
gcattggtaa ctgtcagacc aagtttactc atataactt tagattgatt taaaacttca 4020
tttttaattt aaaaggatct aggtgaagat cctTTTTgat aatctcatga ccaaaatccc 4080
ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc 4140
ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc 4200
agcggTggtt tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt 4260
cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt 4320


```

caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc 4380
tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa 4440
ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag ccagccttgg agcgaacgac 4500
ctacaccgaa ctgagatacc tacagcgtga gcattgagaa agcgccaacgc ttcccgaagg 4560
gagaaaggcg gacagggtatc cggtaagcgg cagggtcgga acaggagagc gcacgagggg 4620
gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact 4680
tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa 4740
cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt tctttcctgc 4800
gttatccctt gattctgtgg ataaccgtat taccgccttt gaggagctg ataccgctcg 4860
ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcgggaag agcgccaat 4920
acgcaaaccg cctctccccg cgcgttgggc gattcattaa tccagctggc acgacaggtt 4980
tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttacc tcactcatta 5040
ggcaccaccg gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg 5100
tattgttaaa gtgtgtcctt tgtcgatact ggtactaatg ctttaatt 5147

```

```

<210> 2
<211> 5171
<212> DNA
<213> Artificial

```

```

<220>
<223> plasmid pSV.ID circular ds-DNA

```

```

<220>
<221> misc_feature
<222> (444)..(444)
<223> splice donor

```

```

<220>
<221> misc_feature
<222> (1946)..(1946)
<223> start pUC118

```

```

<220>
<221> misc_feature
<222> (1954)..(1954)
<223> linearization linker inserted into HpaI site

```

```

<220>

```

<221> misc_feature
 <222> (529)..(1090)
 <223> DHFR coding region

<220>
 <221> misc_feature
 <222> (1522)..(1522)
 <223> sv40 origin

<400> 2
 ttcgagctcg cccgacattg attattgact agagtcgac gacagctgtg gaatgtgtgt 60
 cagttagggt gtggaaagtc cccaggtcc ccagcaggca gaagtatgca aagcatgcat 120
 ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg 180
 caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gcccatcccg 240
 ccctaactc cgcccagttc cgcccattct cgcgccatg gctgactaat tttttttatt 300
 tatgcagagg ccgaggccgc ctgggctct gagctattcc agaagtagtg aggaggcttt 360
 tttggaggcc taggcttttg caaaaagcta gcttatccgg ccgggaacgg tgcattggaa 420
 cgcggtattcc ccgtgccaa agtgacgtaa gtaccgcta tagagtctat agggccaccc 480
 cttggctcta gagagatata agcctaggat tttatcccc gtgccatcat ggttcgacca 540
 ttgaactgca tcgtcgccgt gtcccaaat atggggattg gcaagaacgg agacctacc 600
 tgccctccgc tcaggaacgc gttcaagtac ttccaaagaa tgaccacaac ctcttcagt 660
 gaaggtaaac agaactctgt gattatgggt aggaaaacct ggttctccat tcctgagaag 720
 aatcgacctt taaaggacag aattaatata gttctcagta gagaactcaa agaaccacca 780
 cgaggagctc attttcttg caaagtttg gatgatgcct taagacttat tgaacaaccg 840
 gaattggcaa gtaaagtaga catggtttg atagtcggag gcagttctgt ttaccaggaa 900
 gccatgaatc aaccaggcca ccttagactc tttgtgacaa ggatcatgca ggaatttgaa 960
 agtgacacgt ttttccaga aattgatttg gggaaatata aacctctccc agaataccca 1020
 ggcgtcctct ctgaggcca ggaggaaaaa ggcacaaagt ataagtttga agtctacgag 1080
 aagaaagact aacaggaaga tgctttcaag ttctctgctc ccctoctaaa gctatgcatt 1140
 tttataagac catgggactt ttgctggctt tagacccctc tggcttcgtt agaacgcggc 1200
 tacaattaat acataacctt atgtatcata cacatagatt taggtgacac tatagaataa 1260
 catccacttt gcctttctct ccacaggtgt cactccaggt caactgcacc tcggttctat 1320
 cgattgaatt ccccggggat cctctagagt cgacctgcag aagcttggcc gccatggccc 1380

aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 1440
aataaagcat ttttttact gcattctagt tgtggtttgt ccaaactcat caatgtatct 1500
tatcatgtct ggatcgatcg ggaattaatt cggcgcagca ccatggcctg aaataacctc 1560
tgaaagagga acttggttag gtaccttctg aggcggaaag aaccagctgt ggaatgtgtg 1620
tcagttaggg tgtggaaagt cccaggctc cccagcaggc agaagtatgc aaagcatgca 1680
totcaattag tcagcaacca ggtgtgaaa gtccccaggc tccccagcag gcagaagtat 1740
gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc cgcccatccc 1800
gcccctaact ccgcccagtt ccgcccattc tccgcccatt ggctgactaa ttttttttat 1860
ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 1920
ttttggaggc ctaggctttt gcaaaaagct gttacctoga gcggccgctt aattaaggcg 1980
cgccatttaa atcctgcagg taacagcttg gactggccg tcgttttaca acgtcgtgac 2040
tgggaaaacc ctggcgttac ccaacttaat cgcttgag caccatcccc cttcgccagc 2100
tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc aacagttgcg tagcctgaat 2160
ggcgaatggc gcctgatgcg gtattttctc cttacgcac tgtgcggtat ttcacaccgc 2220
atacgtcaaa gcaaccatag tacgcgcctt gttagcggcg attaagcgcg gcgggtgtgg 2280
tggttacgcg cagcgtgacc gctacacttg ccagcgcctt agcgcgcgt cctttcgtt 2340
tcttccctc ctttctgcc acgttgcgcg gctttccccg tcaagctcta aatcgggggc 2400
tccttttagg gttccgattt agtgctttac ggcacctoga ccccaaaaaa cttgatttgg 2460
gtgatggttc acgtagtgg ccacgcctt gatagacggt tttcgcctt ttgacgttgg 2520
agtccacgtt ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct 2580
cgggctattc ttttgattta taagggattt tgccgatttc ggcctatttg ttaaaaaatg 2640
agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaattttat 2700
ggtgcactct cagtacaatc tgctctgatg ccgcatagtt aagccaactc cgctatcgct 2760
acgtgactgg gtcattggctg cgccccgaca cccgccaaca cccgtgacg cgccctgacg 2820
ggcttgtctg ctcccggcat ccgcttacag acaagctgtg accgtctccg ggagctgcat 2880
gtgtcagagg ttttcaccgt catcaccgaa acgcgcgagg cagtattctt gaagacgaaa 2940
gggcctcgtg atacgcctat ttttataggt taatgtcatg ataataatgg tttcttagac 3000
gtcagggtggc acttttcggg gaaatgtgcg cggaaccctt atttgtttat ttttctaaat 3060

acattcaa atgtatccgc tcatgagaca ataaccotga taaatgcttc aataatattg 3120
aaaaagga agtatgagta ttcaacattt ccgtgtcgcc cttattccct tttttgcggc 3180
attttgcctt cctgtttttg ctcaaccaga aacgctgggt aaagtaaaag atgctgaaga 3240
tcagttgggt gcacgagtgg gttacatcga actggatctc aacagcggtg agatccttga 3300
gagttttcgc cccgaagaac gttttccaat gatgagcaat tttaaagtgc tgctatgtgg 3360
cgcggtatta tcccgatgat acgcccggca agagcaactc ggtcgccgca tacactattc 3420
tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac 3480
agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact 3540
tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca 3600
tgtaactcgc cttgatcggt gggaaccgga gctgaatgaa gccataccaa acgacgagcg 3660
tgacaccacg atgccagcag caatggcaac aacgttgccg aaactattaa ctggcgaaact 3720
acttactcta gcttcccggc aacaattaat agactggatg gaggcggata aagttgcagg 3780
accacttctg cgctcggccc ttccggctgg ctggtttatt gctgataaat ctggagccgg 3840
tgagcgtggg tctcgcggta tcattgcagc actggggcca gatggttaag cctcccgtat 3900
cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc 3960
tgagataggt gcctcactga ttaagcattg gtaactgtca gaccaagttt actcatatat 4020
acttttagatt gatttaaaac ttcatTTTTA atttaaagg atctaggtga agatcctttt 4080
tgataatctc atgacaaaaa tcccttaacg tgagttttcg ttccactgag cgtcagaccc 4140
cgtagaaaag atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt 4200
gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg ccggatcaag agctaccaac 4260
tcctttttccg aaggtaactg gcttcagcag agcgcagata ccaaatactg tccttctagt 4320
gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct 4380
gctaatacctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttggg 4440
ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac 4500
acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagcattg 4560
agaaagcgcc acgcttcccg aaggggagaaa ggcggacagg tatccggtaa gcggcagggg 4620
cggaacagga gagcgcacga gggagcttcc agggggaaac gcctggtatc tttatagtcc 4680
tgtcgggttt cgccacctct gaacttgagc tcgatttttg tgatgctcgt cagggggggc 4740
gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttcttggcct tttgctggcc 4800

```

ttttgctcac atgttctttc ctgcgttata ccttgattct gtggataacc gtattaccgc 4860
ctttgagtga gctgataccg ctgcgccgag ccgaacgacc gagcgcagcg agtcagtga 4920
cgaggaagcg gaagagcgcc caatacgcaa accgcctctc cccgcgcggt ggccgattca 4980
ttaatccagc tggcagcaca ggtttcccg a ctggaaagcg ggcagtgagc gcaacgcaat 5040
taatgtgagt tacctcactc attaggcacc ccaggcttta cactttatgc ttccggctcg 5100
tatgttgtgt ggaattgtga gcggataaca atttcacaca ggaaacagct atgaccatga 5160
ttacgaatta a 5171

```

```

<210> 3
<211> 5712
<212> DNA
<213> Artificial

```

```

<220>
<223> plasmid pSV.IPD circular ds-DNA

```

```

<220>
<221> misc_feature
<222> (444)..(444)
<223> splice donor

```

```

<220>
<221> misc_feature
<222> (479)..(479)
<223> start PUR coding

```

```

<220>
<221> misc_feature
<222> (1079)..(1643)
<223> DHFR coding region

```

```

<400> 3
ttcgagctcg cccgacattg attattgact agagtcgac gacagctgtg gaatgtgtgt 60
cagttagggt gtggaaagtc ccaggctcc ccagcaggca gaagtatgca aagcatgcat 120
ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcaggc cagaagtatg 180
caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gcccatcccg 240
cccctaactc cgcccagttc cgcccattct ccgccccatg gctgactaat tttttttatt 300
tatgcagagg ccgaggccgc ctggcctct gagctattcc agaagtagtg aggaggcttt 360
tttgagggcc taggcttttg caaaaagcta gttatccgg ccgggaacgg tgcattggaa 420

```

cgcggtattcc ccgtgccaaag agtgacgtaa gtaccgccta tagagcgact agtccacat 480
 gaccgagtag aagccacagg tgcgcctcgc caccgcgcac gacgtcccgc gggccgtacg 540
 caccctcgcc gccgcgttcg ccgactaccc cgcacgcgc cacaccgtag acccggaaccg 600
 ccacatcgag cgggtcaccg agctgcaaga actcttcctc acgcgcgtcg ggctcgacat 660
 cggcaagggtg tgggtcgcgg acgacggcgc cgcgggtggcg gtctggacca cgccggagag 720
 cgctgaagcg ggggcgggtg tgcgcgagat cggccgcgc atggccgagt tgagcggttc 780
 ccggctggcc gcgcagcaac agatggaagg cctcctggcg ccgcaccggc ccaaggagcc 840
 cgcgtggttc ctggccaccg tcggcgtctc gcccgaccac cagggcaagg gtctggcgag 900
 cgccgtcgtg ctcccggag tggaggcggc cgagcgcgc ggggtgccc ccttcctgga 960
 gacctcgcg ccccgcaacc tccccttcta cgagcggctc ggcttcaccg tcaccgccga 1020
 cgctcgagtgc ccgaaggacc gcgcgacctg gtgcatgacc cgcaagccc gtgccaacat 1080
 ggttcgacca ttgaactgca tcgtcgccgt gtccaaaat atggggattg gcaagaacgg 1140
 agacctaccc tgccctcgc tcaggaacgc gttcaagtac ttccaaagaa tgaccacaac 1200
 ctcttcagtg gaaggtaaac agaactcgtt gattatgggt aggaaaacct ggttctccat 1260
 tcctgagaag aatcgacctt taaaggacag aattaatata gttctcagta gagaactcaa 1320
 agaaccacca cgaggagctc attttcttgc caaagtttg gatgatgcct taagacttat 1380
 tgaacaaccg gaattggcaa gtaaagtaga catggtttg atagtcggag gcagttctgt 1440
 ttaccaggaa gccatgaatc aaccaggcca ccttagactc tttgtgacaa ggatcatgca 1500
 ggaatttgaa agtgacacgt ttttccaga aattgatttg gggaaatata aacctctccc 1560
 agaataccca ggcgtcctct ctgaggcca ggaggaaaaa ggcatcaagt ataagttga 1620
 agtctacgag aagaaagact aacgttaact gctcccctcc taaagctatg catttttata 1680
 agaccatggg acttttgctg gctttagatc cccttggtt cgttagaacg cagctacaat 1740
 taatacataa ccttatgtat catacacata cgatttaggt gacactatag ataacatcca 1800
 ctttgccctt ctctccacag gtgtccactc ccaggcca ctgcacctcg gttctatcga 1860
 ttgaattccc cggggtcct ctagagtoga cctgcagaag cttcgatggc cgccatggcc 1920
 caacttggtt attgcagctt ataatggta caaataaagc aatagcatca caaatttcac 1980
 aaataaagca tttttttcac tgcattctag ttgtggtttg tccaaactca tcaatgtatc 2040
 ttatcatgtc tggatcgatc gggaattaat tcggcgcagc accatggcct gaaataacct 2100
 ctgaaagagg aacttggtta ggtaccttct gaggcgaaa gaaccagctg tggaatgtgt 2160

gtcagttagg gtgtggaaag tccccaggct cccagcagg cagaagtatg caaagcatgc 2220
atctcaatta gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta 2280
tgcaaagcat gcatctcaat tagtcagcaa ccatagtccc gccctaact ccgcccattcc 2340
cgccccctaac tccgcccagt tccgcccatt ctccgcccc tggttgacta atttttttta 2400
tttatgcaga ggccgaggcc gctcggcct ctgagctatt ccagaagtag tgaggaggct 2460
tttttgagg cctaggcttt tgcaaaaagc tgttacctcg agcggccgct taattaaggc 2520
gcgccattta aatcctgcag gtaacagctt ggcactggcc gtcgttttac aacgtcgtga 2580
ctgggaaaac cctggcgcta cccaactta tcgccttgca gcacatcccc ccttcgccag 2640
ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc gtagcctgaa 2700
tgggcaatgg cgctgatgc ggtattttct ccttacgcat ctgtgcggta tttcacaccg 2760
catacgtcaa agcaaccata gtacgcgcc tgtagcggcg cattaagcgc ggcgggtgtg 2820
gtggttacgc gcagcgtgac cgctacactt gccagcgccc tagcgcgcgc tcctttcgct 2880
ttcttccctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct aaatcgggg 2940
ctccctttag ggttccgatt tagtgcttta cggcacctcg acccaaaaa acttgatttg 3000
ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg ttttcgccc tttgacgttg 3060
gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc 3120
tcgggctatt cttttgattt ataagggtt ttgccgattt cggcctattg gttaaaaaat 3180
gagctgattt aacaaaaatt taacgcgaat ttaacaaaa tattaacgtt tacaatttta 3240
tggtgcactc tcagtacaat ctgctctgat gccgcatagt taagccaact ccgctatcgc 3300
tacgtgactg ggtcatggct gcgccccgac acccgccaac acccgctgac gcgccctgac 3360
gggcttgtct gctccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca 3420
tgtgtcagag gttttaccg tcatcacga aacgcgcgag gcagtattct tgaagacgaa 3480
agggcctcgt gatagccta tttttatagg ttaatgtcat gataataatg gtttcttaga 3540
cgtcaggtag cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa 3600
tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt 3660
gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc tttttgcgg 3720
cattttgcct tcctgttttt gctcaccag aaacgctggg gaaagtaaaa gatgctgaag 3780
atcagttggg tgcacgagtg gggtacatcg aactggatct caacagcggg aagatccttg 3840

agagttttcg cccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg 3900
gcgcgggtatt atcccgatgat gacgcgggc aagagcaact cggtcgccc atacactatt 3960
ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga 4020
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaaacttac 4080
ttctgacaac gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggagc 4140
atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc 4200
gtgacaccac gatgccagca gcaatggcaa caacgttgcg caaactatta actggcgaac 4260
tacttactct agcttcccg caacaattaa tagactggat ggaggcggat aaagttgcag 4320
gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg 4380
gtgagcgtgg gtctcgggt atcattgcag cactggggcc agatggtaag ccctcccgta 4440
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg 4500
ctgagatagg tgccctactg attaagcatt ggtaactgtc agaccaagtt tactcatata 4560
tacttttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt 4620
ttgataatct catgacaaa atcccttaac gtgagttttc gttccactga gcgtcagacc 4680
ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct 4740
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa 4800
ctctttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag 4860
tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc 4920
tgctaatect gttaccagtg gctgtgccca gtggcgataa gtcgtgtctt accgggttg 4980
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca 5040
cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagcatt 5100
gagaaagcgc cacgcttccc gaaggagaa aggcggacag gtatccggtg agcggcaggg 5160
tcggaacagg agagcgacg agggagcttc cagggggaaa cgcttggtat ctttatagtc 5220
ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc 5280
ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc 5340
cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg 5400
cctttgagtg agctgatacc gctcgcgcga gccgaacgac cgagcgcagc gagtcagtga 5460
gcgaggaagc ggaagagcgc ccaatacga aaccgcctct cccgcgcgt tggccgattc 5520
attaatccag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa 5580

ttaatgtgag ttacctcact cattaggcac cccaggcttt acactttatg cttccggctc 5640
 gtatgttgtg tggaattgtg agcggataac aatttcacac aggaacagc tatgaccatg 5700
 attacgaatt aa 5712

<210> 4
 <211> 12514
 <212> DNA
 <213> Artificial

<220>
 <223> plasmid pSV.IPD.2C4 circular ds-DNA

<220>
 <221> misc_feature
 <222> (444)..(444)
 <223> splice donor

<220>
 <221> misc_feature
 <222> (479)..(479)
 <223> start PUR coding

<220>
 <221> misc_feature
 <222> (1079)..(1643)
 <223> DHFR coding region

<220>
 <221> misc_feature
 <222> (1883)..(1883)
 <223> start 2C4 HC coding

<220>
 <221> misc_feature
 <222> (4154)..(4154)
 <223> start LC coding

<400> 4
 ttcgagctcg cccgacattg attattgact agagtcgac gacagctgtg gaatgtgtgt 60
 cagttagggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 120
 ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg 180
 caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gcccatcccg 240
 ccctaactc cgcccagttc cgcccattct cgcgcccatg gctgactaat tttttttatt 300

tatgcagagg ccgaggccgc ctggcctct gagctattcc agaagtagtg aggaggcttt 360
tttgagggcc taggcttttg caaaaagcta gcttatccgg ccgggaacgg tgcattggaa 420
cgcgattcc ccgtgccaa agtgacgtaa gtaccgccta tagagcgact agtccaccat 480
gaccgagtag aagcccacgg tggcctcgc caccgcgac gacgtccgc gggccgtacg 540
caccctcgcc gccgcgttc cgactacce cgccacgcgc cacaccgtag acccggaccg 600
ccacatcgag cgggtcaccg agctgcaaga actcttctc acgcgcgtcg ggctcgacat 660
cggcaagggtg tgggtcgcgg acgacggcgc cgcggtggcg gtctggacca cgccggagag 720
cgtcgaagcg ggggcggtgt tcgcgcgagat cgcccgcggc atggccgagt tgagcggttc 780
ccggctggcc gcgcagcaac agatggaagg cctcctggcg ccgcaccggc ccaaggagcc 840
cgcggtggtc ctggccaccg tcggcgtctc gcccgaccac cagggaagg gtctgggcag 900
cgccgtcgtg cccccggag tggaggcggc cgagcgcgc ggggtgcccg ccttctgga 960
gacctccgc ccccgcaacc tccccctta cgagcgctc ggcttcaccg tcaccgccga 1020
cgtcgagtc ccgaaggacc gcgcgacctg gtgcatgacc cgcaagcccg gtgccaacat 1080
ggttcgacca ttgaactgca tcgtcgccgt gtccaaaat atggggattg gcaagaacgg 1140
agacctacc tgccctcgc tcaggaacgc gttcaagtac ttccaaagaa tgaccacaac 1200
ctcttcagt gaaggtaa acgaatctgt gattatgggt aggaacacct ggttctccat 1260
tcctgagaag aatcgacct taaaggacag aattaatata gttctcagta gagaactcaa 1320
agaaccacca cgaggagctc atttcttgc caaaagttg gatgatgcct taagacttat 1380
tgaacaaccg gaattggcaa gtaaagtaga catggtttg atagtcggag gcagttctgt 1440
ttaccaggaa gccatgaatc aaccaggcca ccttagactc tttgtgacaa ggatcatgca 1500
ggaatttgaa agtgacacgt tttcccaga aattgattg gggaaatata aacctctccc 1560
agaataccca ggcgtcctc ctgaggtcca ggaggaaaa ggcacatcaat ataagtttga 1620
agtctacgag aagaaagact aacgttaact gctccctcc taaagctatg ctttttata 1680
agaccatggg acttttgctg gcttttagat cccttggtt ogttagaacg cagctacaat 1740
taatacataa ccttatgtat catacacata cgatttaggt gacactatag aataacatcc 1800
actttgcctt tctctccaca ggtgtccact ccaggtcca actgcacctc ggttctatcg 1860
attgaattcc accatgggat ggtcatgtat catcctttt ctagtagcaa ctgcaactgg 1920
agtacattca gaagttcagc tgggtggagt tggcggtggc ctggtgcagc cagggggctc 1980

actccgtttg tcctgtgcag cttctggctt caccttcacc gactatacca tggactgggt 2040
 ccgtcaggcc ccgggtaagg gcctggaatg ggttgcatg gttaatccta acagtggcgg 2100
 ctctatctat aaccagcgct tcaagggccg ttctactctg agtggtgaca gatctaataa 2160
 cacattatac ctgcagatga acagcctgcg tgctgaggac actgccgtct attattgtgc 2220
 tcgtaacctg ggacctctt tctactttga ctactgggt caaggaaccc tggtcaccgt 2280
 ctctcggcc tccaccaagg gcccatcggg cttccccctg gcacctcct ccaagagcac 2340
 ctctgggggc acagcgcccc tgggctgcct ggtcaaggac tacttccccg aaccgggtgac 2400
 ggtgtcgtgg aactcaggcg ccctgaccag cggcgtgcac accttccccg ctgtcctaca 2460
 gtcctcagga ctctactccc tcagcagcgt ggtgactgtg ccctctagca gcttgggcac 2520
 ccagacctac atctgcaacg tgaatcaca gccagcaac accaagggtg acaagaaagt 2580
 tgagcccaaa tcttgtgaca aaactcacac atgccaccg tgcccagcac ctgaactcct 2640
 ggggggaccg tcagtcttcc tcttcccccc aaaacccaag gacacctca tgatctcccg 2700
 gacctctgag gtcacatgcg tgggtgtgga cgtgagccac gaagacctg aggtcaagtt 2760
 caactggtac gtggacggcg tggaggtgca taatgccaag acaaagccgc gggaggagca 2820
 gtacaacagc acgtaccggg tggtcagcgt cctcacgcgc ctgcaccagg actgggtgaa 2880
 tggcaaggag tacaagtga aggtctccaa caaagccctc ccagccccca togagaaaac 2940
 catctccaaa gccaaagggc agccccgaga accacagggt tacacctgc ccccatcccg 3000
 ggaagagatg accaagaacc aggtcagcct gacctgcctg gtcaaaggct tctatcccag 3060
 cgacatcgcc gtggagtggg agagcaatgg gcagccggag acaactaca agaccacgcc 3120
 tcccgtgctg gactccgacg gctccttctt cctctacagc aagctcacgc tggacaagag 3180
 cagggtggcag caggggaacg tcttctcatg ctccgtgatg catgaggctc tgcacaacca 3240
 ctacacgcag aagagcctct ccctgtctcc gggtaaataa gtgcgacggc cctagagtcg 3300
 acctgcagaa gcttcgatgg ccgccatggc ccaacttgtt tattgcagct tataatgggt 3360
 acaaataaag caatagcatc acaaatttca caaataaagc atttttttca ctgcattcta 3420
 gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctggatcggg aattaattcg 3480
 gcgcagcacc atggcctgaa ataacctctg aaagaggaaac ttggttaggt accttctgag 3540
 gcggaaagaa ccagctgtgg aatgtgtgtc agttagggtg tggaaagtcc ccaggctccc 3600
 cagcaggcag aagtatgcaa agcatgcac tcaattagtc agcaaccagg tgtggaaagt 3660
 cccaggtctc ccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca 3720

tagtcccgcc cctaactccg cccatcccg cctaactcc gccagttcc gccattctc 3780
cgcccatgg ctgactaatt ttttttattt atgcagaggc cgaggccgcc tcggcctctg 3840
agctattcca gaagtagtga ggaggctttt ttggaggact aggcttttgc aaaaagctag 3900
cttatccggc cggaacggt gcattggaac gcggattccc cgtgccaaga gtcaggtaag 3960
taccgcctat agagtctata ggcccacccc cttggcttcg ttagaacgcg gctacaatta 4020
atacataacc ttttgatog atcctactga cactgacatc cactttttct ttttctccac 4080
agggtgtcac tcccagggtc aactgcaact cggttcgcga agctagcttg ggctgcatcg 4140
attgaattcc accatgggat ggtcatgtat catccttttt ctagtagcaa ctgcaactgg 4200
agtacattca gatatccaga tgaccagtc cccgagctcc ctgtccgcct ctgtgggcga 4260
tagggtcacc atcacctgca aggccagtc ggatgtgtct attgggtgcg cctggtatca 4320
acagaaacca ggaaaagtc cgaaactact gatttactcg gcttctacc gataactgg 4380
agtcccttct cgcttctctg gatccggttc tgggacggat ttcactctga ccatcagcag 4440
tctgcagcca gaagacttcg caacttatta ctgtcaaca tattatattt atccttacac 4500
gtttggacag ggtaccaagg tggagatcaa acgaactgtg gctgcacat ctgtcttcat 4560
cttcccgcca tctgatgagc agttgaaatc tggaactgct tctgttgtgt gcctgtgaa 4620
taacttctat cccagagagg ccaaagtaca gtggaagggtg gataacgcc tccaatcggg 4680
taactccag gagagtgtca cagagcagga cagcaaggac agcacctaca gcctcagcag 4740
caccctgacg ctgagcaaag cagactacga gaaacacaaa gtctacgcct gcgaagtcac 4800
ccatcagggc ctgagctcgc ccgtcacaaa gagcttcaac aggggagagt gttaagcttc 4860
gatggccgcc atggcccaac ttgtttattg cagcttataa tggttacaaa taaagcaata 4920
gcatcacaaa tttcacaaat aaagcatttt tttactgca ttctagtgtt ggtttgtcca 4980
aactcatcaa tgtatcttat catgtctgga tcgggaatta attcggcgca gcaccatggc 5040
ctgaaataag tttaaaccct ctgaaagagg aacttgggtta ggtaccgact agtagcaagg 5100
tcgccacgca caagatcaat attaacaatc agtcatctct ctttagcaat aaaaagggtga 5160
aaaattacat tttaaaaatg acaccataga cgatgtatga aaataatcta cttggaaata 5220
aatctaggca aagaagtgc agactgttac ccagaaaact tacaatttgt aaatgagagg 5280
ttagtgaaga tttaaatgaa tgaagatcta aataaactta taaattgtga gagaaattaa 5340
tgaatgtcta agttaatgca gaaacggaga gacatactat attcatgaac taaaagactt 5400

aatattgtga aggtatactt tcttttcaca taaatttgta gtcaatatgt tcaccccaaa 5460
aaagctgttt gttaacttgt caacctcatt tcaaatgta tatagaaagc ccaaagacaa 5520
taacaaaaat attcttgtag aacaaaatgg gaaagaatgt tccactaaat atcaagattt 5580
agagcaaagc atgagatgtg tggggataga cagtgaggct gataaaatag agtagagctc 5640
agaaacagac ccattgatat atgtaagtga cctatgaaaa aaatatggca ttttacaatg 5700
ggaaaatgat gatctttttc ttttttagaa aaacagggaa atatatttat atgtaaaaaa 5760
taaaagggaa cccatatgtc ataccatata cacaaaaaaa ttccagtga ttataagtct 5820
aaatggagaa ggcaaaaactt taaatctttt agaaaataat atagaagcat gccatcatga 5880
cttcagtgtg gagaaaaatt tcttatgact caaagtccta accacaaaga aaagattgtt 5940
aattagattg catgaatatt aagacttatt tttaaaatta aaaaaccatt aagaaaagtc 6000
agggcataga atgacagaaa atatttcaa caccacagta aagagaattg taatatgcag 6060
attataaaaa gaagtcttac aaatcagtaa aaaataaaac tagacaaaaa tttgaacaga 6120
tgaaagagaa actctaaata atcattacac atgagaaact caatctcaga aatcagagaa 6180
ctatcattgc atatacacta aattagagaa atattaaaag gctaagtaac atctgtggca 6240
atattgatgg tatataacct tgatatgatg tgatgagaac agtactttac cccatgggct 6300
tcctcccaaa acccttacc cagtataaat catgacaaat atacttttaa aaccattacc 6360
ctatatctaa ccagtactcc tcaaaactgt caaggtcatc aaaaataaga aaagtctgag 6420
gaactgtcaa aactaagagg aaccaagga gacatgagaa ttatatgtaa tgtggcattc 6480
tgaatgagat ccgagaacag aaaaagaaca gtagctaaaa aactaatgaa atataaataa 6540
agtttgaact ttagtttttt ttaaaaaaga gtagcattaa cacggcaaag tcattttcat 6600
atttttcttg aacattaagt acaagtctat aattaaaaat tttttaaatg tagtctggaa 6660
cattgccaga aacagaagta cagcagctat ctgtgctgtc gcctaactat ccatagctga 6720
ttggtctaaa atgagatata tcaacgctcc tccatgtttt ttgttttctt tttaaatgaa 6780
aaactttatt ttttaagagg agtttcaggt tcatagcaaa attgagagga aggtacattc 6840
aagctgagga agttttctc tattcctagt ttactgagag attgcatcat gaatgggtgt 6900
taaatTTTgt caaatgcttt ttctgtgtct atcaatatga ccatgtgatt ttcttcttta 6960
acctgttgat gggacaaatt acgttaattg attttcaaac gttgaaccac ccttacatat 7020
ctggaataaa ttctacttgg ttgtggtgta tattttttga tacattcttg gattcttttt 7080
gctaataatt ttgtgaaaat gtttgtatct ttgttcatga gagatattgg tctgtgtttt 7140

tcttttcttg taatgtcatt ttctagttcc ggtattaagg taatgctggc ctagttgaat 7200
gatttaggaa gtattccctc tgcttctgtc ttctgaggta ccgcggccgc ccgtcgtttt 7260
acaacgtcgt gactgggaaa accctggcgt taoccaaactt aatgccttg cagcacatcc 7320
ccctttcgcc agctggcgta atagcgaaga ggccgcgacc gatcgccctt cccaacagtt 7380
gcgcagcctg aatggcgaat ggcgccctgat gcggtatttt ctcccttacgc atctgtgcgg 7440
tatttcacac cgcatacgtc aaagcaacca tagtacgcgc cctgtagcgg cgcattaagc 7500
gcggcggttg tggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc 7560
gtccttttcg ctttcttccc ttcttttctc gccacgttcg ccggctttcc ccgtcaagct 7620
ctaaatcggg ggctcccttt agggttccga tttagtgtt tacggcacct cgaccccaaa 7680
aaacttgatt tgggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc 7740
cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca 7800
ctcaacccta tctcgggcta ttcttttgat ttataaggga ttttgccgat ttccggcctat 7860
tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attttaacaa aatattaacg 7920
tttacaattt tatggtgcac tctcagtaca atctgtcttg atgccgcata gttaagccag 7980
ccccgacacc cgccaacacc cgctgacgcg ccctgacggg cttgtctgct ccgcgcaccc 8040
gcttacagac aagctgtgac cgtctccggg agctgcatgt gtcagagggt ttcaccgtca 8100
tcaccgaaac gcgcgagaga cgaaagggcc tcgtgatacg cctattttta taggttaatg 8160
tcatgataat aatggtttct tagacgtcag gtggcacttt tcggggaaat gtgcgcggaa 8220
cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac 8280
cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg 8340
tcgcccttat tccctttttt gcggcatttt gccttctgt ttttgctcac ccagaaacgc 8400
tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg 8460
atctcaacag cggttaagatc cttgagagtt ttcccccga agaacgtttt ccaatgatga 8520
gcacttttaa agttctgcta tgtggcgcggt tattatcccg tattgacgcc gggcaagagc 8580
aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag 8640
aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtgcc ataaccatga 8700
gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagetaaccg 8760
cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga 8820

atgaagccat accaaacgac gagcgtgaca ccacgatgcc ttagcaatg gcaacaacgt 8880
 tgcgcaaact attaaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact 8940
 ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttcg gctggctggt 9000
 ttattgctga taaatctgga gccggtgagc gtgggtctcg cgtatcatt gcagcactgg 9060
 gccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta 9120
 tggatgaacg aaatagacag atcgtgaga taggtgcctc actgattaag cattggtaac 9180
 tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta 9240
 aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt 9300
 tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt 9360
 ttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt 9420
 gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc 9480
 agataccaaa tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg 9540
 tagcacgcc tacatactc gctctgctaa tctgttacc agtggctgct gccagtggcg 9600
 ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg 9660
 cgggctgaac ggggggttcg tgcacacagc ccagcttga gcaacgacc tacaccgaac 9720
 tgagatacct acagcgtgag ctatgagaaa gcgccacgt tccgaaggg agaaggcgg 9780
 acaggtatcc ggtaagcggc agggctcgaa caggagagcg cagagggag cttccagggg 9840
 gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat 9900
 ttttgtgatg ctctcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt 9960
 tacggttctt ggcttttgc tggccttttg ctacatggt ctttctgctg ttatccctg 10020
 attctgtgga taaccgtatt accgcctttg agtgagctga tacgctcgc cgcagccgaa 10080
 cgaccgagcg cagcgagtca gtgagcgagg aagcgggaaga gcccgcgggc aaggctcgcca 10140
 cgcacaagat caatattaac aatcagtcatt ctctctttag caataaaaag gtgaaaaatt 10200
 acattttaaa aatgacacca tagacgatgt atgaaaataa tctacttgga aataaatcta 10260
 ggcaaagaag tgcaagactg ttaccagaa aacttacaaa ttgtaatga gaggttagtg 10320
 aagatttaaa tgaatgaaga tctaaataaa cttataaatt gtgagagaaa ttaatgaatg 10380
 tctaagttaa tgcagaaacg gagagacata ctatattcat gaactaaaag acttaatt 10440
 gtgaaggat actttctttt cacataaatt tgtagtcaat atgttcaccc caaaaaagct 10500
 gtttgtaac ttgtcaacct catttcaaaa tgtatataga aagcccaaag acaataacaa 10560

aaatattctt gtagaacaaa atgggaaaga atgttccact aaatatcaag atttagagca 10620
aagcatgaga tgtgtgggga tagacagtga ggctgataaa atagagtaga gctcagaaac 10680
agaccattg atatatgtaa gtgacctatg aaaaaaatat ggcatTTTtac aatgggaaaa 10740
tgatgatctt tttctTTTT agaaaaacag ggaaatatat ttatatgtaa aaaataaaag 10800
ggaaccata tgcatacca tacacacaaa aaaattccag tgaattataa gtctaaatgg 10860
agaaggcaaa actttaaatc ttttagaaaa taatatagaa gcatgccatc atgaCTTcag 10920
tgtagagaaa aatttcttat gactcaaagt cctaaccaca aagaaaagat tgTTaattag 10980
attgcatgaa tattaagact tattttttaa attaaaaaac cattaagaaa agtcaggcca 11040
tagaatgaca gaaaatattt gcaacacccc agtaaagaga attgtaatat gcagattata 11100
aaaagaagtc ttacaaatca gtaaaaaata aaactagaca aaaatttgaa cagatgaaag 11160
agaaactcta aataatcatt acacatgaga aactcaatct cagaaatcag agaactatca 11220
ttgcatatac actaaattag agaaatatta aaaggctaag taacatctgt ggcaatattg 11280
atggtatata acctgatat gatgtgatga gaacagtact ttacCCcatg ggcttCctcc 11340
ccaaaccctt accccagtat aaatcatgac aaatatactt taaaaccat taccctatat 11400
ctaaccagta ctCctcaaaa ctgtcaaggt catcaaaaat aagaaaagtc tgaggaactg 11460
tcaaaactaa gaggaaccca aggagacatg agaattatat gtaatgtggc attctgaatg 11520
agatcccaga acagaaaaag aacagtagct aaaaaactaa tgaaatataa ataaagtTtg 11580
aactttagtt ttttttaaaa aagagtagca ttaacacggc aaagtcattt tcatattttt 11640
cttgaacatt aagtacaagt ctataattaa aaatttttta aatgtagtct ggaacattgc 11700
cagaaacaga agtacagcag ctatctgtgc tgtcgctaa ctatccatag ctgattggtc 11760
taaaatgaga tacatcaacg ctCctccatg tttttgttt tctttttaa tgaaaaactt 11820
tattttttaa gaggagtttc aggttcatag caaaattgag aggaaggtac attcaagctg 11880
aggaagtttt cctctattcc tagtttactg agagattgca tcatgaatgg gtgttaaatt 11940
ttgtcaaagc cttttctgt gtctatcaat atgacatgt gattttcttc ttTaaacctgt 12000
tgatgggaca aattacgtta attgattttc aaacgttgaa ccacccttac atatctggaa 12060
taaattctac ttggttgTgg tgtatatttt ttgatacatt cttggattct ttttgctaatt 12120
atTTtgTga aaatgtttgt atctttgttc atgagagata ttggtctgtt gttttctttt 12180
cttgTaatgt catTTtctag ttccggtatt aaggtaatgc tggcctagtt gaatgattta 12240

ggaagtattc cctctgcttc tgtcttctga agcggaagag cgccaatac gcaaaccgcc 12300
tctccccg cggtggcga ttcattaatg cagctggcac gacaggtttc ccgactggaa 12360
agcgggcagt gagcgcaacg caattaatgt gagttagctc actcattagg caccacaggc 12420
tttacacttt atgcttccgg ctcgatatgt gtgtggaatt gtgagcgat aacaatttca 12480
cacaggaaac agctatgaca tgattacgaa ttaa 12514

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.